Econometric Analysis of Panel Data Models with Multifactor Error Structures

Karabiyik, Hande; Palm, Franz C.; Urbain, Jean Pierre

published in
Annual Review of Economics
2019

DOI (link to publisher)
10.1146/annurev-economics-063016-104338

document version
Publisher's PDF, also known as Version of record

ANNUAL REVIEWS

Annual Review of Economics

Econometric Analysis of Panel Data Models with Multifactor Error Structures

Hande Karabiyik,[1] Franz C. Palm,[2] and Jean-Pierre Urbain[2,*]

[1]Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands

[2]Department of Quantitative Economics, Maastricht University, 6200 MD Maastricht, The Netherlands; email: f.palm@maastrichtuniversity.nl

Annu. Rev. Econ. 2019. 11:495–522

The *Annual Review of Economics* is online at economics.annualreviews.org

https://doi.org/10.1146/annurev-economics-063016-104338

*Prof. Jean-Pierre Urbain passed away on October 1, 2016.

JEL codes: B23, C01, C23, C33, C55, F17

Keywords

panel data, cross-sectional dependence, factor-augmented panel regression, common correlated effects, principal components, stationary panels, nonstationary panels

Abstract

Economic panel data often exhibit cross-sectional dependence, even after conditioning on appropriate explanatory variables. Two approaches to modeling cross-sectional dependence in economic panel data are often used: the spatial dependence approach, which explains cross-sectional dependence in terms of distance among units, and the residual multifactor approach, which explains cross-sectional dependence by common factors that affect individuals to a different extent. This article reviews the theory on estimation and statistical inference for stationary and nonstationary panel data with cross-sectional dependence, particularly for models with a multifactor error structure. Tests and diagnostics for testing for unit roots, slope homogeneity, cointegration, and the number of factors are provided. We discuss issues such as estimating common factors, dealing with parameter plethora in practice, testing for structural stability and nonlinearity, and dealing with model and parameter uncertainty. Finally, we address issues related to the use of these economic panel models.

1. INTRODUCTION

Increased availability of panel data sets in economics with large cross-section and time dimensions has been paralleled by a similar supply of models and methods for studying panel data in economics and business. In areas such as finance, macroeconomics, and international and regional studies, it is now customary to work with panel data to back policy decisions by empirical research using time series and panel data. Governments of countries and regions and international and national organizations such as business firms and banks rely on the policy conclusions drawn from empirical panel data studies.

The popularity of panel data studies is also due to important recent developments in the scientific literature on analyzing panel data and to the availability of computer power and software to carry out sophisticated computations.

The early literature on panel data methods assumes that, conditional on some individual characteristics, different entities such as countries, firms, and individuals are independent of each other. However, analyses of panel data sets with large time series and cross-section dimensions have shown that the individual units in many data sets are interdependent. For instance, in cross-country growth analysis, this type of interdependence is highly likely to reflect the interconnectedness of countries through history, geography, and trade relations (Eberhardt & Teal 2011). Relevant theoretical research has shown that the presence of this interdependence may have serious consequences, depending on the nature of dependence, if not accounted for. This insight led to a new branch of theoretical research that develops models and methods to account for cross-sectional dependence in economic panel data.

Two main approaches deal with cross-sectional dependence in large panels, namely, the spatial dependence approach and the residual multifactor approach. The former assumes that dependence across cross-section units can be characterized by a spatial process that represents the distance among units. The spatial approach has been pioneered by Anselin (1988), and since then, a vast literature has been developed (for a recent survey, see Lee & Yu 2010). One of the limitations of using this approach is that it typically does not allow for slope heterogeneity across panel units. Another limitation is that the approach requires a priori knowledge of a distance metric that might be difficult to come up with in many applications. The residual multifactor approach assumes that dependence among cross-section units is caused by a fixed and usually small number of common factors that affect all individual units to a different extent. This assumption is easily justified considering the existence of global shocks that affect all units to a different extent in, for example, macroeconomic or financial data. Examples of such situations abound and are discussed briefly in this review. One advantage of this approach is that it does not require prior knowledge about the ordering of the cross-section units.

This article provides a review of the methods that use the residual multifactor approach for analyzing large panel data sets with cross-sectional dependence. We focus on estimation and inference problems when all variables involved are stationary and exogenous, when the model is dynamic, and when the variables are nonstationary and possibly cointegrated. We review the literature on panel unit root tests for panels with cross-sectional dependence and on tests for cointegration. Furthermore, we provide a discussion of issues arising in panel modeling, such as how to estimate the factors, how to deal with parameter plethora in practice, the implications of nonlinearity and structural changes for panel data modeling, and model and parameter uncertainty.

The remainder of this review is organized as follows. In Section 2, we describe the residual multifactor approach to cross-sectional dependence. In Section 3, we discuss estimation of large panel data models by distinguishing between different settings that vary depending on the static versus dynamic and stationary versus nonstationary nature of the variables. Section 4 contains a review of inferential methods and diagnostic tests, such as unit root tests and tests for cointegration, that are applicable to large panel data sets with cross-sectional dependence.

In Section 5, we discuss the issues associated with panel data modeling, such as nonlinearity and structural changes. Section 6 contains a short discussion on the use of panel data models for prediction and policy making. Section 7 concludes.

2. CROSS-SECTIONAL DEPENDENCE IN PANELS

Economic, financial, and climate time series data across cross-section units such as countries, regions, firms, households, and individuals are likely to be affected by some global shocks. These shocks affect each unit with different intensities, depending on the intrinsic properties of each cross-section unit. For instance, output growth and CO_2 emissions across multiple countries have the tendency to be affected by the same technology shocks, natural disasters, and global financial crises. The extent to which a certain country is affected by these shocks depends on factors such as the capital intensity of its production process, its openness measures, and its trade volume. Bailey et al. (2016) investigate the cross-sectional correlation among stock returns. Motivated by the capital asset pricing model (Ross 1976), they assume that there are market factors that create correlation among all assets in a financial market. Ertur & Musolesi (2017) investigate the influence of technological knowledge spillovers on total factor productivity. They acknowledge the presence of unobserved common factors, such as aggregate technological shocks or oil price shocks, that may have different effects on total factor productivity across countries. They suggest that heterogeneous effects of these factors may result from country-specific technological constraints.

In general, the existence of global factors creates an interdependence among the cross-section units. To deal with this interdependence, the literature has adopted an approach that is called the unobserved common factors or interactive fixed effects approach. This way of modeling cross-sectional dependence has become increasingly popular over the past decades. An example is the global vector autoregressive (VAR) model proposed by, among others, Pesaran et al. (2004) to model regional economic dependencies in a global macroeconomic model. In a first step, small-scale country-specific VAR models that are augmented with foreign variables are estimated. In a second step, these models are stacked as one large global VAR model. This is discussed further in Section 5.2. In this section, we introduce a simple unobserved common factor model and make a distinction between strong and weak cross-sectional dependence.

Consider the process $z_{i,t}$ that is observable over $t = 1, \ldots, T$ and $i = 1, \ldots, N$. We assume that each $z_{i,t}$ exhibits statistical dependence across $i = 1, \ldots, N$, which can be modeled by using the unobserved common factor approach. We write the model for $z_{i,t}$ as

$$z_{i,t} = \boldsymbol{\lambda}_i'\mathbf{f}_t + e_{i,t}, \qquad 1.$$

where $\mathbf{f}_t$ is an $r \times 1$ vector of unobserved common shocks, $\boldsymbol{\lambda}_i$ is an $r \times 1$ vector of unit-specific unknown factor loadings, and $e_{i,t}$ are idiosyncratic time- and unit-specific shocks. This is the formal representation of an unobserved common factor model and is used in the literature to model the cross-sectional dependence between $z_{i,t}$ and $z_{j,t}$ for $i \neq j$ (for a review of econometric analysis of large factor models, see Bai & Wang 2016). The analysis of cross-sectionally dependent panel data models depends on the assumptions regarding how the factors, $\mathbf{f}_t$, enter the models and on the properties of $\boldsymbol{\lambda}_i$ and $\mathbf{f}_t$.

2.1. Weak and Strong Cross-Sectional Dependence

In the literature, the degree of cross-sectional dependence in panel data models is categorized into three main groups: strong, weak, and semistrong or semiweak cross-sectional dependence (Bai 2003, 2009; Bailey et al. 2016; Chudik & Pesaran 2015b; Chudik et al. 2011; Sarafidis & Wansbeek 2012). The degree of cross-sectional dependence determines the effects of the dependence on the processes and on the estimators and, in turn, influences the ways in which this cross-sectional dependence should be handled.

Let $\mathrm{Var}(\mathbf{z}_t) = \mathbf{\Sigma}_{z,t}$, denote the covariance matrix of $\mathbf{z}_t$, where $\mathbf{z}_t = (z_{1,t}, z_{2,t}, \ldots, z_{N,t})'$ and where $\mathbf{\Sigma}_{z,t}$ is an $N \times N$ symmetric, nonnegative definite matrix. The definition for weak cross-sectional dependence adopted by Bai (2003, 2009) is based on $\mathbf{\Sigma}_{z,t}$. Inspired by the assumptions of the approximate factor model of Chamberlain & Rothschild (1983), we define weak cross-sectional correlation by

$$\frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{N}|\sigma_{ij,t}| \leq K_t < \infty \qquad 2.$$

for all N, where $\sigma_{ij,t}$ is the element at the intersection of the ith column and the jth row of $\mathbf{\Sigma}_{z,t}$. This imposes a bound on the sum of the absolute value pairwise correlations between $z_{i,t}$ and $z_{j,t}$.

Chudik et al. (2011) and Bailey et al. (2016) define the concepts of weak and strong cross-sectional dependence based on the limiting behavior of the weighted cross-sectional average of the process as $N \to \infty$ for a given $t \in T$. They define $\bar{z}_{w,t} = \sum_{i=1}^{N} w_i z_{i,t}$ as the weighted cross-sectional average of $z_{i,t}$ at time t. In this case, the vector of weights, defined as $\mathbf{w}_t = (w_{1,t}, w_{2,t}, \ldots, w_{N,t})'$, is assumed to satisfy certain granularity conditions for all t such that, for each element of this vector,

$$\|w_t\| = \sqrt{w_t' w_t} = O(N^{-1/2}) \quad \text{and} \quad \frac{w_{j,t}}{\|w_t\|} = O(N^{-1/2}) \quad \text{uniformly in } j \in N.$$

Chudik et al. (2011) give the condition for weak cross-sectional dependence as

$$\lim_{N\to\infty} \mathrm{Var}(\bar{z}_{w,t}) = 0 \qquad 3.$$

and the condition for strong cross-sectional dependence as

$$\lim_{N\to\infty} \mathrm{Var}(\bar{z}_{w,t}) \geq K > 0. \qquad 4.$$

The condition in Equation 3 is not specific about the rate of the convergence, and for a process, there is a range of possibilities between being a strongly dependent process and an independent process. To see this, we write

$$\lim_{N\to\infty} N^{\alpha}\,\mathrm{Var}(\bar{z}_{w,t}) < K < \infty. \qquad 5.$$

In this case, the range of possibilities mentioned above is determined by the exponent α. For example, if $\alpha = 1$, then the process is independent; if $\alpha = 0$, then the process satisfies the condition in Equation 4 and is strongly dependent. The condition in Equation 3 holds for $0 < \alpha \leq 1$ given that Equation 5 holds. Bailey et al. (2016) show how to estimate α under certain conditions. Besides the weak and strong cross-sectional dependence defined by Chudik et al. (2011), another type of dependence occurs when a particular unit has a dominant influence on the rest of the units. This is analyzed by Chudik & Pesaran (2013). It is also possible to assume that the dependence is caused by both spatial effects and common factors. This situation is investigated by Pesaran & Tosetti (2011).

2.2. Weak and Strong Factors

Besides the categorization of the degrees of cross-sectional dependence, a categorization regarding the strength of the factors is also considered. In particular, Chudik et al. (2011) define the set of strong factors as the factors whose loadings satisfy

$$\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} |\lambda_{i,s}| = K > 0 \qquad 6.$$

and the set of weak factors as the factors whose loadings satisfy the condition

$$\lim_{N\to\infty} \sum_{i=1}^{N} |\lambda_{i,s}| = K < \infty, \qquad 7.$$

where $\lambda_{i,s}$ is the factor loading that corresponds to factor $f_{s,t}$. Furthermore, if

$$\lim_{N\to\infty} N^{-\kappa} \sum_{i=1}^{N} |\lambda_{i,s}| = K < \infty \qquad 8.$$

with $0 < \kappa < \frac{1}{2}$ ($\frac{1}{2} \leq \kappa < 1$), then the factor $f_{s,t}$ belongs to the set of semiweak (semistrong) factors. The relation between strong (weak) factors and strong (weak) cross-sectional dependence is analyzed by Chudik et al. (2011). They show that, if one of the factors in Equation 1 is strong, then $z_{i,t}$ is strongly cross-sectionally dependent.

2.3. Testing for Cross-Sectional Dependence

Consequences of cross-sectional dependence might be severe. Thus, there is a need to test for the cross-sectional dependence of the data before analyzing a panel data set. The literature on testing for cross-sectional dependence is reviewed by Chudik & Pesaran (2015b) and Sarafidis & Wansbeek (2012). One strand of the literature contains tests that are based on the pairwise correlations of the regression errors (see, for example, Breusch & Pagan 1980; Pesaran 2004, 2015; Pesaran et al. 2008; Sarafidis et al. 2009).

3. ESTIMATING PANEL DATA MODELS WITH CROSS-SECTIONAL DEPENDENCE

In this section, we introduce the linear panel model with a multifactor error structure. We provide a detailed discussion of how unobserved factors may influence the asymptotic properties of estimators of model parameters. We discuss two of the most popular approaches used in estimating such models, namely the principal components (PC) and the common correlated effects (CCE) approaches. We provide a review of these approaches within static stationary, dynamic stationary, and nonstationary frameworks. We provide a discussion on the advantages and limitations of these approaches. We give examples of empirical applications that consider these approaches.

Suppose that we observe the scalar $y_{i,t}$ and an n_x-vector of individual specific regressors, $\mathbf{x}_{i,t}$, over T time periods and on N cross-section units, where $t = 1, \ldots, T$ and $i = 1, \ldots, N$. The scalar $y_{i,t}$ is assumed to follow the general model

$$y_{i,t} = \boldsymbol{\gamma}_i' \mathbf{d}_t + \boldsymbol{\beta}_i' \mathbf{x}_{i,t} + u_{i,t}, \qquad 9.$$

where $\mathbf{d}_t$ is a n_d-vector of observed common effects that can also include deterministic components, $\boldsymbol{\gamma}_i$ is the n_d-vector of coefficients of the observed common effects, and $\boldsymbol{\beta}_i$ is the n_x-vector of heterogeneous slope coefficients. In this case, the n_x-vector $\mathbf{x}_{i,t}$ can include exogenous variables and/or lagged values of $y_{i,t}$. Finally, $u_{i,t}$ is the error term that exhibits dependence across cross-section units, i.e., $E(u_{i,t}u_{j,t}) \neq 0$ for some $i \neq j$ with $i, j \in \{1, \ldots, N\}$. A multifactor error structure that is used to model this dependence is given by

$$u_{i,t} = \boldsymbol{\lambda}_i' \mathbf{f}_t^y + \varepsilon_{i,t}, \qquad 10.$$

where $\mathbf{f}_t^y$ is an r_y-vector of unobserved common factors, $\boldsymbol{\lambda}_i$ is the r_y-vector of factor loadings that represents the extent to which each cross-section unit is affected by the unobserved common factors, and $\varepsilon_{i,t}$ is the idiosyncratic error of the cross-section unit i at time t. Suppose that $\mathbf{x}_{i,t}$ follows the data generating process (DGP)

$$\mathbf{x}_{i,t} = \boldsymbol{\Gamma}_i \mathbf{d}_t + \boldsymbol{\Lambda}_i \mathbf{f}_t^x + \boldsymbol{\nu}_{i,t}, \qquad 11.$$

where $\mathbf{f}_t^x$ is an r_x-vector of unobserved common factors, $\boldsymbol{\Lambda}_i$ is the $n_x \times r_x$ matrix of factor loadings, and $\boldsymbol{\nu}_{i,t}$ is a n_x-vector of idiosyncratic errors.

If $\boldsymbol{\Lambda}_i \mathbf{f}_t^x$ is independent of $\boldsymbol{\lambda}_i' \mathbf{f}_t^y$, then the inclusion of $\mathbf{f}_t^y$ in Equation 10 does not affect the consistency of the least squares type estimators of $\boldsymbol{\beta}_i$, in general. However, if the DGP of $\mathbf{x}_{i,t}$ is a nondegenerate function of $\mathbf{f}_t^y$, for instance, if the elements of $\mathbf{f}_t^y$ are also the elements of $\mathbf{f}_t^x$, then this consistency might be lost. This is, of course, due to the fact that the intersection of $\mathbf{f}_t^x$ and $\mathbf{f}_t^y$ creates dependence between the errors and the regressors in Equation 9, leading to endogeneity. Note that, in this case, we allow for $\mathbf{f}_t^y$ to be different from $\mathbf{f}_t^x$. Now consider $\mathbf{f}_t^y, \mathbf{f}_t^x$ as sets of factors. Let $\mathbf{f}_t^{y\cup x} \equiv \mathbf{f}_t^y \cup \mathbf{f}_t^x$, $\mathbf{f}_t^{y\cap x} \equiv \mathbf{f}_t^y \cap \mathbf{f}_t^x$, $\mathbf{f}_t^{y\setminus x} \equiv \mathbf{f}_t^y \setminus \mathbf{f}_t^x$, and $\mathbf{f}_t^{x\setminus y} \equiv \mathbf{f}_t^x \setminus \mathbf{f}_t^y$. The factors that would lead to a loss of consistency of the ordinary least squares (OLS) estimator of $\boldsymbol{\beta}_i$ are the ones in $\mathbf{f}_t^{y\cap x}$. The cardinality of this set is the number of effective factors, and this is the number that needs to be taken into account. Note that the elements in $\mathbf{f}_t^{y\setminus x}$ and $\mathbf{f}_t^{x\setminus y}$ might be ignored if the aim is to obtain a consistent estimator of $\boldsymbol{\beta}_i$. Ignoring the effective factors will lead to inconsistency and problems with standard inferential methods.

One of the most popular approaches to estimating Equation 9 is the PC approach. This approach uses PC analysis to extract the factors. The idea of using the PC approach in panel data models to account for cross-sectional dependence goes back to Coakley et al. (2002). Another approach uses cross-sectional averages to find estimators for the space spanned by the factors and is called the CCE approach. The CCE approach was first proposed by Pesaran (2006).[1] For both methods, an estimator for the factors is obtained by using the observed data in a first step. In a second step, the regression model is augmented by the estimated factors, and the augmented model is estimated by using the appropriate estimation method to assure consistency and valid inference.

[1]Garcia-Ferrer et al. (1987) use the cross-sectional median instead of the cross-sectional mean to estimate the common effect of a panel of nine countries in a more robust way.

One of the distinctions that we make when reviewing the literature on the analysis of the framework given in Equations 9–11 is among stationary static panels, stationary dynamic panels, and nonstationary panels. Furthermore, distinguishing among different frameworks regarding the heterogeneity of $\boldsymbol{\beta}_i$, $\boldsymbol{\lambda}_i$, and $\boldsymbol{\Lambda}_i$ is also important. In the heterogeneous slope coefficient case, $\boldsymbol{\beta}_i$ is allowed to vary across cross-section units with a limited variation, and in the homogeneous slope coefficient case, $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all $i = 1, \ldots, N$. The choice of the estimators, the derivation, and the rates of convergence of the estimators relies on the heterogeneity and homogeneity properties of the parameters of interest. Phillips & Sul (2003) propose a Hausman-type test for slope homogeneity for stationary $AR(1)$ models with cross-sectional dependence that works well for panels with small N and large T. Pesaran & Yamagata (2008) propose a test for slope homogeneity for large panels; however, their test does not deal with the serial correlation and heteroskedasticity of the errors. Blomquist & Westerlund (2013) propose a generalized test that accommodates both serial correlation and heteroskedasticity of the errors.

3.1. Large Heterogeneous Static Stationary Panels with a Multifactor Error Structure

The multifactor error structure approach in a static model with stationary panel data has been used in numerous estimation exercises in empirical studies. For instance, by using this approach, Bertoli & Moraga (2013) estimate the determinants of migration rates in the context of a general individual utility maximization model. Taking into account the multivariate error structure, they find a smaller effect of GDP per capita and a larger effect of migration policies on bilateral migration rates than the effects suggested by estimation strategies that assume cross-sectional independence.

Consider the framework in Equations 9–11, and let $\mathbf{f}_t = \mathbf{f}_t^y = \mathbf{f}_t^x$, $r = r_y = r_x$. Pesaran (2006) develops the CCE approach to estimate $\boldsymbol{\beta}_i$ and its mean, $\boldsymbol{\beta}$. This approach suggests using the weighted cross-sectional averages of the observed variables as an approximation for the linear combinations of the factors. In particular, by letting $\mathbf{z}_{i,t} = (y_{i,t}, \mathbf{x}_{i,t}')'$, one can define the weighted cross-sectional averages as $\bar{\mathbf{z}}_{w,t} = \sum_{i=1}^{N} w_i \mathbf{z}_{i,t}$, where $\{w_i\}$ are the weights that satisfy some very general granularity conditions.[2] We can write the augmented regression model as

$$y_{i,t} = \boldsymbol{\gamma}_i'\mathbf{d}_t + \boldsymbol{\beta}_i'\mathbf{x}_{i,t} + \boldsymbol{\delta}_i'\bar{\mathbf{z}}_{w,t} + \varepsilon_{i,t}^*. \qquad 12.$$

Note the difference between $\boldsymbol{\lambda}_i$ in Equation 10 and $\boldsymbol{\delta}_i$ in Equation 12; this is because, when we replace the true factors with the weighted cross-sectional averages, the loadings also change and become a rotation of the original loadings. This implies that $\boldsymbol{\lambda}_i$ is not directly estimable by the regression in Equation 12. Additionally, note that the error terms of the models in Equations 10 and 12 are not the same. The difference stems from the fact that $\varepsilon_{i,t}^*$ includes the factor approximation error as well as the regression error. This can be formally written as $\varepsilon_{i,t}^* = \varepsilon_{i,t} + \boldsymbol{\delta}_i'[\boldsymbol{\delta}_i(\boldsymbol{\delta}_i'\boldsymbol{\delta}_i)^{-1}\boldsymbol{\lambda}_i'\mathbf{f}_t - \bar{\mathbf{z}}_{w,t}]$, where the second term is the factor approximation error. Pesaran (2006) shows that, under certain conditions, this second term is asymptotically negligible as $N \to \infty$.

It is convenient to write the estimators in matrix notation. For this reason, we define $\bar{\mathbf{h}}_{w,t} = (\mathbf{d}_t', \bar{\mathbf{z}}_{w,t}')'$ as the $(n_d + n_x + 1)$-vector of common regressors, and then we define $\bar{\mathbf{H}}_w = (\bar{\mathbf{h}}_{w,1}, \ldots, \bar{\mathbf{h}}_{w,T})'$, $\mathbf{X}_i = (\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,T})'$ and $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,T})'$, which have dimensions

[2]A simplified case can involve considering equal weights $w_i = 1/N$ for $i = 1, \ldots, N$. Then $\bar{\mathbf{z}}_{w,t}$ will be the simple cross-sectional averages.

$T \times (n_d + n_x + 1)$, $T \times n_x$, and $T \times 1$, respectively. The OLS-type estimator of $\boldsymbol{\beta}_i$ can be written as

$$\widehat{\boldsymbol{\beta}}_{\mathrm{CCE},i} = (\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{y}_i, \tag{13}$$

where $\mathbf{M}_{\overline{H}} = \mathbf{I}_T - \overline{\mathbf{H}}_w(\overline{\mathbf{H}}_w'\overline{\mathbf{H}}_w)^{-1}\overline{\mathbf{H}}_w'$ is the usual orthogonal projection matrix. For the mean of the individual coefficients, Pesaran (2006) considers two different pooling methods. The first is called the pooled CCE estimator and can be written as

$$\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}} = \left(\sum_{i=1}^{N} w_i\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{X}_i\right)^{-1}\sum_{i=1}^{N} w_i\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{y}_i. \tag{14}$$

Note that the pooling weights are equal to the weights used to aggregate the cross-sectional dimension to obtain the weighted averages as an approximation for the factors. The second is called the mean group CCE estimator and can be written as

$$\widehat{\boldsymbol{\beta}}_{\mathrm{CCEMG}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\boldsymbol{\beta}}_{\mathrm{CCE},i} = \frac{1}{N}\sum_{i=1}^{N}\left[(\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{M}_{\overline{H}}\mathbf{y}_i\right]. \tag{15}$$

In this case, we note an important point. Asymptotic properties of the estimators $\widehat{\boldsymbol{\beta}}_{\mathrm{CCE},i}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ depend on a certain condition known as the rank condition. To see what the rank condition requires, recall that $\mathbf{z}_{i,t} = (y_{i,t}, \mathbf{x}_{i,t}')'$, and let

$$\mathbf{z}_{i,t} = \mathbf{B}_i'\mathbf{d}_t + \mathbf{C}_i'\mathbf{f}_t + \boldsymbol{\epsilon}_{i,t}, \tag{16}$$

where $\mathbf{C}_i = (\boldsymbol{\lambda}_i + \boldsymbol{\Lambda}_i'\boldsymbol{\beta}_i, \boldsymbol{\Lambda}_i')$ and $\mathbf{B}_i$ and $\boldsymbol{\epsilon}_{i,t}$ can be defined similarly. The rank condition is imposed on the weighted average of $\mathbf{C}_i$, such that $\mathrm{rank}\,(\overline{\mathbf{C}}_w) = r \leq n_x + 1$, where $\overline{\mathbf{C}}_w = \sum_{i=1}^{N} w_i\mathbf{C}_i$. This rank condition is primarily about the number of factors in the model and the number of observed variables. It requires the number of unobserved factors to be not greater than the number of observed panel data variables. In general, the CCE method works better when this condition holds. Note that the number of effective factors argument holds here. If $\mathbf{f}_t^y \neq \mathbf{f}_t^x$, then the rank condition changes to rank $(\mathbf{f}_t^y \cap \mathbf{f}_t^x) \leq n_x + 1$ (Juodis et al. 2018).

Pesaran (2006) shows that, under certain conditions, $\widehat{\boldsymbol{\beta}}_{\mathrm{CCE},i}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEMG}}$ are consistent and asymptotically normally distributed as $N, T \to \infty$. The rank condition and the relative growth rates of N and T play an important role in these results. One finding of Pesaran (2006) is that $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ is asymptotically biased if, for example, $T/N \to \kappa$, where $0 < \kappa < \infty$. This might be an issue for panels with similar N and T dimensions. An analytical expression for this bias is derived by Westerlund & Urbain (2015). It is important to note that no assumptions are made on the correlation between $\boldsymbol{\lambda}_i$ and $\boldsymbol{\Lambda}_i$ by Pesaran (2006). However, for the heterogeneous slope coefficient case, Westerlund & Urbain (2013) show that $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ is inconsistent if the loadings are correlated when the rank condition is not satisfied.

Pesaran (2006) assumes only the presence of strong cross-sectional dependence. Chudik et al. (2011) consider the estimation of the slope coefficient $\boldsymbol{\beta}$ when $\varepsilon_{i,t}$ and $\boldsymbol{v}_{i,t}$ are affected by infinitely many weak factors that satisfy Equation 8. They show that, in this case, $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEMG}}$ are consistent and asymptotically normally distributed if the number of weak factors grows at a certain rate bounded by N and κ as $N, T \to \infty$. Pesaran & Tosetti (2011) investigate the case with both unobserved common factors and spatial error correlation in the errors $u_{i,t}$ and find that allowing for spatial correlation in the errors does not alter the results obtained by Pesaran (2006). A more robust version of the CCE method can be to use the median of the cross-sectional data instead of the average. This method was first considered by Garcia-Ferrer et al. (1987). Since then, there have been no crucial developments in this method.

Building on the work of Coakley et al. (2002), Bai (2009) considers the estimation of Equation 9 by using the PC approach. The PC approach of Coakley et al. (2006) yields consistent estimates of the slope parameters only when the regressors are strongly exogenous for the factors, in particular, when $\mathbf{\Lambda}_i = 0$. Bai (2009) proposes using an iterative approach to obtain consistent estimates even when the regressors are correlated with the factors. Bai (2009) sets $\boldsymbol{\gamma}_i = 0$, $\mathbf{\Gamma}_i = 0$ and $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all $i = 1, \ldots, N$ in Equations 9–11. The estimator for $\boldsymbol{\beta}$ is obtained by solving

$$\widehat{\boldsymbol{\beta}}_{\mathrm{PC}} = \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{F_{\mathrm{PC}}} \mathbf{X}_i\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{F_{\mathrm{PC}}} \mathbf{y}_i \qquad 17.$$

and

$$\frac{1}{NT} \sum_{i=1}^{N} \left(\mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{PC}}\right) \left(\mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{PC}}\right)' \widehat{\mathbf{F}}_{\mathrm{PC}} = \widehat{\mathbf{F}}_{\mathrm{PC}} \widehat{\mathbf{V}}, \qquad 18.$$

where $\mathbf{M}_{F_{\mathrm{PC}}} = \mathbf{I}_T - \widehat{\mathbf{F}}_{\mathrm{PC}} (\widehat{\mathbf{F}}_{\mathrm{PC}}' \widehat{\mathbf{F}}_{\mathrm{PC}})^{-1} \widehat{\mathbf{F}}_{\mathrm{PC}}'$, where the $T \times r$ matrix $\widehat{\mathbf{F}}_{\mathrm{PC}}$ consists of the first r eigenvectors times $\sqrt{T}$ of the matrix $\frac{1}{NT} \sum_{i=1}^{N} (\mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{PC}})(\mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{\mathrm{PC}})'$, and where the $r \times r$ matrix $\widehat{\mathbf{V}}$ is a diagonal matrix that has the r largest eigenvalues of the same matrix arranged in a decreasing order. The implementation can be carried out by starting with an initial value for $\widehat{\boldsymbol{\beta}}_{\mathrm{PC}}$. For instance, one can use the least squares estimator from regression $y_{i,t}$ on $\mathbf{x}_{i,t}$ to obtain $\widehat{\mathbf{F}}_{\mathrm{PC}}$ and $\widehat{\mathbf{V}}$ in the second equation and then use $\widehat{\mathbf{F}}_{\mathrm{PC}}$ in the first equation to obtain the updated $\widehat{\boldsymbol{\beta}}_{\mathrm{PC}}$. Bai (2009) shows that the $\widehat{\boldsymbol{\beta}}_{\mathrm{PC}}$ that is obtained after sufficient iterations is a consistent estimator of $\boldsymbol{\beta}$ as $N, T \to \infty$ jointly. Furthermore, the asymptotic distribution of $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{\mathrm{PC}} - \boldsymbol{\beta})$ is normal and centered around zero in two scenarios. The first assumes that there is no serial correlation and that there is heteroskedasticity in the errors. In this case, the asymptotic distribution of $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{\mathrm{PC}} - \boldsymbol{\beta})$ is normal and centered around zero if $T/N \to 0$ as $N, T \to \infty$. The second is when there is no cross-section correlation and heteroskedasticity in the errors. In this case, the asymptotic distribution of $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{\mathrm{PC}} - \boldsymbol{\beta})$ is normal and centered around zero if $N/T \to 0$ as $N, T \to \infty$. In both cases, we see that $\widehat{\boldsymbol{\beta}}_{\mathrm{PC}}$ is asymptotically biased if $N/T \to \kappa$, where $0 < \kappa < \infty$. Bai (2009) provides an expression for the asymptotic bias, considers the bias-corrected estimator when $N/T \to \kappa$ as $N, T \to \infty$, and shows that the bias-corrected estimator is consistent when $T/N^2 \to 0$ and $N/T^2 \to 0$ as $N, T \to \infty$.

The PC approach assumes that the number of factors, r, is known. To implement this method, one needs to obtain a consistent estimator of the number of factors. Numerous works propose a method to estimate the number of factors in a factor model (see, for example, Amengual & Watson 2007; Bai & Ng 2002, 2007).

An important advantage of the CCE method over the PC approach is that the CCE method does not require a priori knowledge of the number of unobserved factors. However, the rank condition, which is restricting the number of factors, plays a role in the validity of some of the results provided by Pesaran (2006). Firstly, the CCE estimator of the individual slope coefficient is inconsistent when the rank condition is not satisfied. Secondly, in a special case where the slope coefficient is homogeneous, for consistent estimation, the rank condition has to be satisfied. Thus, one needs to check the validity of the rank condition, and this requires knowledge of the number of factors. The rank assumption can be considered as a drawback of the CCE approach, as it limits the minimum number of regressors that should be used in the model. To alleviate this, Karabiyik et al. (2019) propose using multiple weighted cross-sectional averages of the observed variables to augment the model to be used in the estimation. Karabiyik et al. (2017) address a problem with the CCE approach that appears in the empirically relevant case when the number of factors is strictly less than the number of observations used in estimation. More precisely, they address the problem that the use of too many observables causes the second moment matrix of the estimated factors to become asymptotically singular.

Another advantage of the CCE over the PC approach is that it does not require any iterations. Thus, it is computationally simpler than the PC approach. However, the PC approach allows the estimation of the factors and factor loadings up to a rotation. Thus, if one is interested in the factors and their loadings, then the PC approach should be preferred.

Both Pesaran (2006) and Bai (2009) assume that the regressors $\mathbf{x}_{i,t}$ are correlated only with $\mathbf{f}_t$, and they assume stochastic independence between $\mathbf{x}_{i,t}$ and $\varepsilon_{i,t}$. Harding & Lamarche (2011) consider the case where $\mathbf{x}_{i,t}$ is allowed to be correlated with $\varepsilon_{i,t}$. They show that an instrumental variables (IV)–type estimator obtained from the model augmented by the cross-sectional averages of the observed variables and some instruments yields consistent estimates of the slope coefficients.

Westerlund & Urbain (2015) compare the asymptotic properties of $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{PC}}$ under the assumption of slope homogeneity. They find that as $N, T \to \infty$, with $\sqrt{T}/N \to 0$ and $\sqrt{N}/T \to 0$, both estimators are $\sqrt{NT}$ consistent and asymptotically normally distributed. If $T/N \to \kappa$, where κ is a finite, nonzero constant, then both estimators are asymptotically biased. They provide theoretical expressions for these biases and show on which factors they depend. They conclude that the biases depend on the properties of the data, such as the correlation between the estimated factors and model errors. One important conclusion of the work of Westerlund & Urbain (2015) is that, whenever the two dimensions, namely N and T, are close to each other, a bias correction is recommended regardless of whether one uses the PC or the CCE approach. Karabiyik et al. (2017) show that, when $r < n_x + 1$, the expression for the bias of $\widehat{\boldsymbol{\beta}}_{\mathrm{CCEP}}$ is different from that in the case with $r = n_x + 1$.

Sarafidis & Wansbeek (2012) compare the approaches developed by Bai (2009) and Pesaran (2006) in a small Monte Carlo study. They show that the CCE approach breaks down when the rank condition is not satisfied and $\boldsymbol{\lambda}_i$ and $\boldsymbol{\Lambda}_i$ are correlated. In all other cases, the CCE approach outperforms the iterative PC method. This finding is theoretically confirmed by Westerlund & Urbain (2013). Another finding in the literature suggests that the estimation of the number of factors can substantially affect the small sample performance of the PC approach (Pesaran & Kapetanios 2005).

3.2. Dynamic Stationary Panel Data Models with a Multifactor Error Structure

Using a multifactor error structure approach is popular in empirical studies, as, by nature, most economic relationships are dynamic. For instance, Teles & Mussolini (2014) examine how the size of the public debt–to-GDP ratio limits the effects of productive government expenditure on long-term growth by adopting a dynamic panel data model that is augmented by a multifactor error structure.

Chudik & Pesaran (2015a) extend the work of Pesaran (2006) to a heterogeneous panel data setup with lagged dependent variables. They consider

$$y_{i,t} = \gamma_i + \phi_i y_{i,t-1} + \boldsymbol{\beta}'_{0,i}\mathbf{x}_{i,t} + \boldsymbol{\beta}'_{1,i}\mathbf{x}_{i,t-1} + \boldsymbol{\lambda}'_i\mathbf{f}_t + \varepsilon_{i,t}. \qquad 19.$$

Furthermore, they assume that the joint DGP of $\mathbf{x}_{i,t}$ and a number of covariates $\mathbf{q}_{i,t}$ follows

$$\begin{pmatrix} \mathbf{x}_{i,t} \\ \mathbf{q}_{i,t} \end{pmatrix} = \boldsymbol{\gamma}_{c,i} + \boldsymbol{\alpha}_i y_{i,t-1} + \boldsymbol{\Lambda}_i \mathbf{f}_t + \boldsymbol{\nu}_{i,t},$$

where $\mathbf{x}_{i,t}$ and $\mathbf{q}_{i,t}$ are n_x and n_q vectors, respectively. They then consider a CCE-type estimation of $\boldsymbol{\beta}_i$ and $\boldsymbol{\beta} = \mathrm{E}(\boldsymbol{\beta}_i)$ under this setup. Chudik & Pesaran (2015a) show that it is necessary to augment the model with the weighted cross-sectional averages of $\mathbf{z}_{i,t} = (y_{i,t}, \mathbf{x}'_{i,t}, \mathbf{q}'_{i,t})'$ and with their lags up to a certain lag order. Theoretically, one should augment the model with the current and lagged values of the cross-sectional averages with a lag order of infinity to remove the effect of the unobserved factors from the remaining errors completely. However, this is neither possible nor necessary. As long as the coefficients of the lag polynomials decay exponentially as $\ell \to \infty$, the lag order can be truncated. This order is usually a function of T. The augmented regression model should then be

$$y_{i,t} = \gamma_i + \phi_i y_{i,t-1} + \boldsymbol{\beta}'_{0,i}\mathbf{x}_{i,t} + \boldsymbol{\beta}'_{1,i}\mathbf{x}_{i,t-1} + \sum_{\ell=0}^{p_T} \boldsymbol{\delta}_{i,\ell}\overline{\mathbf{z}}_{w,t-\ell} + \varepsilon^*_{i,t}, \qquad 20.$$

where $\overline{\mathbf{z}}_{w,t-\ell} = \sum_{i=1}^{N} w_i \mathbf{z}_{i,t-\ell}$ with $w_{i,t}$ satisfying the granularity conditions, as in the static CCE approach, and p_T is the truncation lag order.

Chudik & Pesaran (2015a) establish that the OLS estimation of Equation 20 yields consistent estimates for the individual specific coefficients, $\boldsymbol{\beta}_i = (\phi_i, \boldsymbol{\beta}'_{0,i}, \boldsymbol{\beta}'_{1,i})'$, if $p_T^3/T \to \kappa$ as $N, T, p_T \to \infty$, where $0 < \kappa < \infty$, and if $n_x + n_q + 1 \geq r$. Additionally, they consider the mean group estimator for $\boldsymbol{\beta}$, defined as in Equation 15, and find that it is consistent under the same conditions that are required for the consistency of the individual specific estimator. However, in this case, it is possible to replace the $n_x + n_q + 1 \geq r$ assumption with the serially uncorrelated factors assumption. Let us denote this estimator with $\boldsymbol{\beta}_{\mathrm{DCCEMG}}$. Chudik & Pesaran (2015a) also show that $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{\mathrm{DCCEMG}} - \boldsymbol{\beta})$ is asymptotically normally distributed if $p_T^3/T \to \kappa_1$, $N/T \to \kappa_2$ as $N, T, p_T \to \infty$, where $0 < \kappa_1, \kappa_2 < \infty$, and again if $n_x + n_q + 1 \geq r$ or $\mathbf{f}_t$ is serially uncorrelated.

Song (2013) considers a similar model to Equation 19 but does not specify a model for $\mathbf{x}_{i,t}$ and allows it to belong to a wider family of DGPs and to include lagged dependent variables as well. Song focuses on the estimation of individual specific slope coefficients by using the iterative method of Bai (2009), by replacing Equation 17 with

$$\widehat{\boldsymbol{\beta}}_{\mathrm{DPC},i} = (\mathbf{X}'_i \mathbf{M}_{F_{\mathrm{PC}}} \mathbf{X}_i)^{-1} \mathbf{X}'_i \mathbf{M}_{F_{\mathrm{PC}}} \mathbf{y}_i.$$

Similarly to Equation 17, this approach assumes that the number of factors is known. The estimator, $\widehat{\boldsymbol{\beta}}_{\mathrm{DPC},i}$, is shown to be $\sqrt{T}$-consistent when $T/N^2 \to 0$ as $N, T \to \infty$, and furthermore, if $\boldsymbol{\nu}_{i,t}, \boldsymbol{\lambda}_i$ and $\varepsilon_{i,t}$ are cross-sectionally independent, then, as $T/N^2 \to 0$ as $N, T \to \infty$, $\sqrt{T}(\widehat{\boldsymbol{\beta}}_{\mathrm{DPC},i} - \boldsymbol{\beta}_i)$ is asymptotically normally distributed. A comparison of the small sample performances of the methods of Chudik & Pesaran (2015a) and Song (2013) can be found in the work of Chudik & Pesaran (2015a).

Moon & Weidner (2017) study linear panel regression models with unobserved common factors and predetermined regressors, such as lagged-dependent variables that exhibit heteroskedasticity. They show that, for the case with predetermined regressors, the least squares estimator proposed by Bai (2009) has two sources of asymptotic bias: the bias that occurs due to heteroskedasticity and cross-section correlation and the bias that occurs due to the predeterminedness of the regressors. As a response, they propose a bias correction method. Moon & Weidner (2015) acknowledge the fact that both Bai (2009) and Moon & Weidner (2017) assume that the number of unobserved common factors is known. They study the estimators proposed in those papers when the number of factors is not known. They find that, as long as the number of factors used in the analysis is at least as large as the true number of factors, the asymptotic results of these papers are not affected.

Chudik & Pesaran (2011) carry out an application of the CCE method to infinite-dimensional VAR models. The model for the ith unit can be written as

$$y_{i,t} = \sum_{j=1}^{N} \phi_{i,j} y_{j,t-1} + u_{i,t}. \tag{21.}$$

In this case, the effects of the lagged values of other units on unit i are parametrically modeled. This leads to the issue of dimensionality, as the number of parameters that are needed to be estimated is N^2 (N parameters for each unit). Chudik & Pesaran (2011) propose a method to deal with such a problem. They assume that each unit has a small finite number of neighbors among the other units and that the coefficients of the nonneighboring units tend to zero as $N \to \infty$. This implies that ignoring the nonneighboring units in the estimation does not affect the consistency of the estimators of the coefficients of the neighboring units. Furthermore, they assume that $y_{i,t}$ exhibits strong cross-sectional dependence, and it follows that $y_{i,t} = \alpha_i + \boldsymbol{\lambda}_i' \mathbf{f}_t + \varepsilon_{i,t}$, where $\varepsilon_{i,t}$ is allowed to be spatially and serially correlated. This assumption extends the model of Pesaran (2006) to a dynamic model where all variables are determined endogenously. Chudik & Pesaran (2011) show that, in such a setup, the estimator for the individual slope coefficients of the neighboring units, which is obtained by the OLS regression of $y_{i,t}$ on the lagged values of the neighboring units, on the current and lagged values of the weighted cross-sectional averages of all the units, and on a constant, is consistent as $N, T \to \infty$. Furthermore, if $T/N \to \kappa$ as $N, T \to \infty$, where $0 < \kappa < \infty$, then this estimator is $\sqrt{N}$ consistent and asymptotically normally distributed.

Chudik & Pesaran (2013) assume that, in Equation 21, there is one dominant unit that affects all the units, and the rest of the units are treated as nonneighbors. In this case, the assumption is that one of the units (for example, the first unit, $j = 1$) is dominant, and its direct or indirect effects on the rest of the system can lead to strong cross-sectional dependence. In particular, the dominant unit acts as a dynamic factor in the models of the remaining units. Then the regression equation for the nondominant units should be

$$y_{i,t} = \phi_i y_{i,t-1} + \sum_{\ell=0}^{p_T} \delta_{i,\ell} y_{1,t-\ell} + \varepsilon_{i,t}, \tag{22.}$$

and the regression equation for the dominant unit should be

$$y_{1,t} = \sum_{\ell=1}^{p_T} \phi_{1,\ell} y_{1,t-\ell} + \varepsilon_{1,t}. \tag{23.}$$

Note how the current and lagged values of the cross-sectional averages are replaced by the current and lagged values of the dominant unit, where p_T is the truncation lag order, as before. Chudik & Pesaran (2013) show that the OLS estimator for ϕ_i obtained by running the corresponding regression, i.e., Equation 22 for nondominant units and Equation 23 for the dominant unit, is consistent if $p_T^3/T \to \kappa$ as $N, T \to \infty$, where $0 < \kappa < \infty$. They further show that, if $p_T^3/T \to \kappa_1$

and $T/N \to \kappa_2$ as $N, T \to \infty$, where $0 < \kappa_1, \kappa_2 < \infty$, then this estimator is $\sqrt{T}$ consistent and asymptotically normally distributed.

In this section, we focus on large panels with large time series and cross-section dimensions where both T and N are assumed to go to infinity jointly. The results provided for large panels do not usually hold for short panels, i.e., panels with small time series dimensions. There is a separate vast literature on analyzing dynamic panel data models with fixed T. We only list some of the important papers, as the main focus of our review is large panels. Especially for dynamic panels, working under the fixed-T assumption brings complications to the analysis and to the results. One of the complications is the bias that arises in such setups, which is called the Nickell bias. Nickell (1981) shows that the within-groups estimator for dynamic panel data regressions is inconsistent for fixed T as $N \to \infty$. The order of this bias is $1/T$ and is shown to be quite sizeable in a small-T context. Solutions proposed to this issue involve using estimation methods such as generalized method of moments (Arellano & Bond 1991, Arellano & Bover 1995, Blundell & Bond 1998), bias corrected within groups methods (Bun & Carree 2005, Hahn & Kuersteiner 2002, Kiviet 1995), and likelihood-based methods (Hsiao et al. 2002, Lancaster 2002, Moreira 2009). The works mentioned above assume cross-sectional independence across units. Phillips & Sul (2007) show that, besides the bias caused by within-group demeaning, unobserved common factors induce additional inconsistency when T is fixed, and this additional inconsistency disappears as $T \to \infty$. For other works that deal with cross-sectional dependence in a dynamic panel data setup with fixed T, the reader is referred to, for example, Phillips & Sul (2003), Sarafidis & Robertson (2009), Choi et al. (2010), and Everaert & De Groote (2016).

3.3. Nonstationary Panels with Multifactor Error Structures

It is common in financial and macroeconomic studies to have nonstationary and possibly cointegrated time series. For panels with unobserved factor structures, the analysis of nonstationary models should take into account the sources of nonstationarities. The nonstationarity might stem from idiosyncratic errors $\varepsilon_{i,t}$ or $\boldsymbol{\nu}_{i,t}$, and/or it might stem from unobserved factors. Furthermore, while investigating and estimating the cointegrating relations between variables conditionally on the fact that they have unit roots, it is important to investigate the role of the nonstationary factors in long-run equilibria. For instance, Eberhardt et al. (2013) estimate the long-run effects of R&D investments on productivity by taking into account the unobserved knowledge spillovers (a term that they use to describe unobserved factors). Holly et al. (2010) investigate the cointegration between real house prices and real per capita incomes, while accounting for possible unobserved common factors.

Westerlund & Urbain (2015) investigate the asymptotic properties of the pooled OLS estimators in the presence of common and idiosyncratic stochastic trends. They find that the estimators' asymptotic behavior depends critically on what is assumed regarding unit root properties of the common and idiosyncratic components of $y_{i,t}$ and $\mathbf{x}_{i,t}$.

Consider again the setup in Equations 9 and 10, and let $\mathbf{f}_t = \mathbf{f}_t^y = \mathbf{f}_t^x$ and $r = r_y = r_x$. Suppose that $\varepsilon_{i,t}$ and $\boldsymbol{\nu}_{i,t}$ are stationary processes, whereas $\mathbf{f}_t$ follows the multivariate unit root process

$$\mathbf{f}_t = \mathbf{f}_{t-1} + \boldsymbol{\upsilon}_t,$$

where $\boldsymbol{\upsilon}_t$ is an r-vector with stationary elements that are distributed independently of the individual specific errors $\varepsilon_{i,s}$ and $\boldsymbol{\nu}_{i,k}$ for all t, s, k, and i. Kapetanios et al. (2011) investigate the properties of the CCE estimation within this framework. They find that, under certain conditions including the rank condition, presence of unit roots in $\mathbf{f}_t$ does not alter the asymptotic properties of $\hat{\boldsymbol{\beta}}_{\text{CCEP}}$,

$\hat{\boldsymbol{\beta}}_{\text{CCEMG}}$ and $\hat{\boldsymbol{\beta}}_{\text{CCE},i}$. They suggest that the facts that $\mathbf{f}_t$ is nonstationary and that $\varepsilon_{i,t}$ is stationary might imply that $y_{i,t}$, $\mathbf{x}_{i,t}$, and $\mathbf{d}_t$ and $\mathbf{f}_t$ are cointegrated. This in turn implies that the convergence rate of $\hat{\boldsymbol{\beta}}_{\text{CCE},i}$ should be T. Surprisingly, they find that the convergence rate of $\hat{\boldsymbol{\beta}}_{\text{CCE},i}$ is still $\sqrt{T}$, instead of the usual cointegrating vector estimators' T-convergence. This is because, in their setup, the defactored $y_{i,t}$ and the defactored $\mathbf{x}_{i,t}$ are actually stationary, and we can see $\boldsymbol{\beta}_{\text{CCE},i}$ as the estimator that is obtained by regressing the former on the latter, which would be a stationary regression.

Westerlund (2018) shows that, except for some basic moment requirements on the factors, the CCE is applicable under very general conditions on $\mathbf{f}_t$. This implies that CCE estimation can accommodate factors with polynomial time trends of any finite degree, seasonal and structural break dummies, factors with unknown heteroskedasticity over time, and factors of any finite order of integration.

Bai & Kao (2006) set $d_t = 1$ for all t and $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i and consider the estimation of $\boldsymbol{\beta}$. They assume that Equation 10 holds. However, they assume that the regressors are generated by a unit root process of order one for all i, such that $\mathbf{x}_{i,t} = \mathbf{x}_{i,t-1} + \boldsymbol{\nu}_{i,t}$, and that the unobserved common factors are stationary. They then propose a two-stage fully modified estimation procedure. Bai et al. (2009) assume a unit root process for the factors similar to that of Kapetanios et al. (2011). In contrast to Kapetanios et al. (2011), they assume that there is no cointegration between $\mathbf{x}_{i,t}$ and $\mathbf{f}_t$, and thus $\boldsymbol{\nu}_{i,t}$ in Equation 11 is assumed to have a unit root. They propose an iterative procedure to estimate $\boldsymbol{\beta}$ and $\mathbf{f}_t$. The iteration procedure is similar to that of Bai (2009) but with a different normalization. In particular, in this case, the normalizing factor in Equation 18 is $1/NT^2$ instead of $1/NT$. This is an adjustment made due to the presence of nonstationarity. These iterations yield the $\hat{\boldsymbol{\beta}}_{\text{Cup}}$ estimator, which is $\sqrt{N}T$ consistent as $(N, T) \to \infty$ sequentially if $\mathbf{x}_{i,t}$ and $\mathbf{f}_t$ are exogeneous. Otherwise, the estimator has an asymptotic bias. Bai et al. (2009) propose a consistent estimator for this bias to correct for it.

An advantage of the procedure of Bai et al. (2009) over that of Kapetanios et al. (2011) is that the iterative procedure yields an estimate of the global stochastic trends. This might be important in some applications, such as that of Eberhardt et al. (2013), where there is an economic intuition behind these trends. The procedure of Bai et al. (2009) assumes that the number of factors is known. Kapetanios et al. (2011) do not require this assumption, but due to the rank condition, knowledge about the number of factors carries some importance.

Ergemen & Velasco (2017) study Equations 9–11 for cases in which both idiosyncratic and unobserved common components are fractionally integrated. Kao et al. (2012) consider the estimation of a long-run relation between a process and some unobserved factors. They assume the existence of a set of variables such that the factors can be estimated by using the PC approach.

4. TESTS AND DIAGNOSTICS FOR PANEL DATA MODELS WITH CROSS-SECTIONAL DEPENDENCE

In this section, we review the literature on panel unit root and panel cointegration tests that allows for cross-sectional dependence. In this literature, unit root and cointegration properties of output, investments, CO_2 emissions, greenhouse gas emissions, health expenditures, and stock returns, among numerous other economic and financial variables, have been investigated using panel data and allowing for cross-sectional dependence (e.g., Baltagi & Moscone 2010, Bond et al. 2010, Eberhardt & Teal 2013, Eberhardt et al. 2013).

4.1. Testing for Unit Roots

It has been shown that univariate unit root tests have low power unless the number of observations is very high (e.g., Dickey & Fuller 1979), especially when the time series exhibits a high level of persistency. Panel unit root tests have increased power compared to univariate unit root tests, which is mostly due to the increased number of observations (Levin et al. 2002). This makes testing for unit roots by using panels attractive and explains the immense growth in the literature on panel unit root tests in the past decades. However, using panels while testing for unit roots leads to certain issues. These issues raise some questions, such as how to deal with the cross-sectional dependence, how to deal with serial correlation, how to deal with time series heteroskedasticity, how to specify the hypotheses to be tested, how to pool the information, how to deal with deterministic trends, and how to deal with structural breaks. In this section, we review the recent literature that addresses these questions.

Consider the general model stated in Equation 9, and let $\mathbf{d}_t = 0$, $x_{it} = y_{i,t-1}$, and $\theta_i = \beta_i - 1$; we have

$$\Delta y_{i,t} = \theta_i y_{i,t-1} + u_{i,t}. \tag{24}$$

A panel unit root test can be constructed by considering the null hypothesis

$$\mathcal{H}_0: \quad \theta_i = 0 \ \text{ for all } i.$$

Earlier literature assumed that $u_{i,t}$ exhibits no or little statistical dependence across i. These tests are categorized as first-generation panel unit root tests. As shown by, for example, O'Connell (1998) and Banerjee et al. (2004), these tests are invalid and have severe size distortions if the cross-sectional independence assumption is not satisfied. This fact has led to the development of second-generation panel unit root tests, which are robust to cross-sectional dependence. Breitung & Pesaran (2008) and Hurlin & Mignon (2007) provide early surveys of the second-generation unit root tests.

4.1.1. Second-generation panel unit root tests. We focus on the three most popular approaches to test for unit roots in cross-sectionally dependent panels. These approaches were developed by Bai & Ng (2004), Pesaran (2007b), and Moon & Perron (2004). The PANIC test of Bai & Ng (2004) can test for unit roots in the unobserved global factors, $\mathbf{f}_t$, and in the idiosyncratic component, $\varepsilon_{i,t}$, separately, whereas the other two tests focus on testing for unit roots in the idiosyncratic components only. The approaches of Bai & Ng (2004) and Moon & Perron (2004) can accommodate multiple common factors, whereas that of Pesaran (2007b) allows for a maximum of one common factor. Bai & Ng (2004) and Moon & Perron (2004) adopt the PC approach, whereas Pesaran (2007b) adopts the CCE approach. Gengenbach et al. (2009) compare the small sample performances of Bai & Ng (2004), Pesaran (2007b), and Moon & Perron (2004) and find that the presence of serial correlation leads to size distortions for almost all tests when the time series dimension is small. The Bai & Ng (2004) test for unit roots in the common factors has low power. The pooled Pesaran (2007b) and the Bai & Ng (2004) test statistics are more powerful than the individual test statistics that they are based on. In the presence of a single factor, the pooled Pesaran (2007b) test statistic has good power and size; however, an increase in the number of factors has negative effects on the power. The Moon & Perron (2004) and Bai & Ng (2004) test statistics have good power properties, with some size distortions, and are not affected by an increase of the number of unobserved factors. Another investigation of the power properties of Bai & Ng (2004) is conducted by Westerlund (2015b), who shows that the (local) power of PANIC is affected negatively by serial correlation and by the heteroskedasticity of the innovations.

Pesaran et al. (2013) extend the work of Pesaran (2007b) to allow for multiple common factors. By ignoring the linear trend and the intercept, the model that they consider reduces to

Equation 24. They assume Equation 10 for the error process. Note that, in Equation 10, the dimension of vector $\mathbf{f}_t$ is $r \times 1$, where $r \geq 1$. Pesaran (2007b) assumes the same structure but restricts $r = 1$ and then suggests using the cross-sectional average of $y_{i,t}$ to find an approximation for the space spanned by this factor. The presence of more than one unobserved factor makes the cross-sectional averages of $y_{i,t}$ insufficient for the approximation of the factors in the system. More observed series are needed to obtain approximations for the space spanned by the unobserved factors. Pesaran et al. (2013) assume that there exists k variables with data generating process $\Delta \mathbf{x}_{i,t} = \mathbf{\Lambda}_i \mathbf{f}_t + \boldsymbol{\nu}_{i,t}$, where $\mathbf{\Lambda}_i$ is a $k \times r$ matrix of factor loadings, and $\boldsymbol{\nu}_{i,t}$ is a k-vector of idiosyncratic errors. They then propose using the cross-sectionally augmented-OLS-based t-test that is based on the regression

$$\Delta y_{i,t} = \theta_i y_{i,t-1} + \mathbf{c}_i' \overline{\mathbf{z}}_{t-1} + \mathbf{h}_i' \overline{\Delta \mathbf{z}}_t + \epsilon_{i,t},$$

where $\overline{\mathbf{z}}_t = (\bar{y}_t, \overline{\mathbf{x}}_t')'$. This is called the cross-sectionally augmented Dickey–Fuller regression. They then construct a usual individual specific t-statistic, say, t_i, to test for $\mathcal{H}_0 : \theta_i = 0$ and show that, under certain conditions, the asymptotic distribution of this statistic is free of nuisance parameters. Furthermore, for the panel unit root test to test for $\mathcal{H}_0 : \theta_i = 0$ for all i against the alternative $\mathcal{H}_A$: $\theta_i < 0$ for $i = 1, 2 \ldots, N_1$ and $\theta_i = 0$ for $i = N_1 + 1, \ldots, N$, where $N_1/N \to \kappa$ and $0 < \kappa \leq 1$ as $N \to \infty$, they consider the cross-sectional average of these statistics, CIPS $= N^{-1} \sum_{i=1}^{N} t_i$ (where CIPS stands for cross-sectionally augmented Im, Pesaran and Shin test). The limiting distribution of this test statistic is the average of the limiting distribution of t_i; thus, it is nuisance parameter free. However, the limiting distribution is not standard. The critical values are given by Pesaran et al. (2013). Contrary to the tests developed by Moon & Perron (2004) and PANIC, this test does not require the estimation of the number of factors. It is valid as long as $k + 1 \geq r$. This test is robust to serial correlation in the errors. To accommodate this, one needs to augment the regression model with $\Delta y_{i,t-1}$, $\overline{\Delta \mathbf{z}}_{t-1}$ and preferably their lags. Noticing that the CIPS statistic is highly nonstandard and requires the use of tabulated critical values, Westerlund & Hosseinkouchack (2016) modify the CIPS test statistic to obtain a statistic that has a standard distribution. Reese & Westerlund (2016) combine the cross-sectional averages method of Pesaran (2006) and the PANIC method and suggest a new unit root test that embodies the advantages of the PANIC and CIPS tests but alleviates their negative aspects.

The tests mentioned above also differ from each other with respect to how they pool the information. The PANIC test is based on the pooled p-values of the individual test statistics. This has the advantage of allowing for more heterogeneity in the autoregressive coefficients. The tests developed by Moon & Perron (2004) are based on the pooled estimator of the autoregressive coefficient. Pesaran et al.'s (2013) CIPS test is based on pooled t-statistics. These types of tests have good power when the autoregressive coefficients are homogeneous across cross-section units.

4.1.2. The incidental trends problem. Related to the issue of deterministic trends, Moon & Phillips (1999) discover that it is difficult to detect unit roots in panels when heterogeneous trends are present. This problem arises because of the presence of an infinite number of nuisance parameters. This phenomenon is called the incidental trends problem. Moon & Perron (2004) show that, when there are incidental trends that are removed by using OLS detrending, their tests suffer from low power. In response to this, Bai & Ng (2010) develop a test that is based on the PANIC methodology and on the pooled estimator of the autoregressive coefficient that does not suffer from the loss of power due to the presence of incidental trends. This approach can be considered as the combination of the PANIC and Moon & Perron (2004) approaches, applying a defactoring inspired by PANIC and a pooling inspired by Moon & Perron (2004). The method that

Bai & Ng (2010) use to defactor the data simultaneously takes care of the detrending of the data. They then obtain the bias-corrected pooled estimator of the autoregressive coefficient and provide test statistics based on this estimator. They show that these test statistics have standard normal distributions asymptotically and have nontrivial power in finite samples. As another solution to the incidental trends problem, Westerlund (2015a) suggests recursive detrending. This approach allows for even nonlinear trends in the DGP. Westerlund shows that the unit root test statistic based on a pooled estimator of the autoregressive coefficient that is obtained by using recursively detrended data is asymptotically normally distributed with mean zero and unit variance. Additionally, Westerlund (2015c) suggests adding covariates to deal with the incidental trends problem. This leads to a reduction in the loss of power that is caused by the presence of incidental trends.

4.1.3. Accounting for heteroskedasticity. The approaches mentioned above assume homoskedasticity of the errors. However, Cavaliere (2005) shows that nonconstant variances affect the size and power properties of the augmented Dickey Fuller–type tests and make them dependent on nuisance parameters. Demetrescu & Hanck (2012) study the asymptotic behavior of panel unit root tests when there is unconditional heteroskedasticity in the innovations. They show that the tests of Moon & Perron (2004), Breitung & Das (2005), and Pesaran (2007b) have size distortions when there is unconditional heteroskedasticity in the time dimension. As a response to this, Demetrescu & Hanck (2012) propose a test that is based on a Cauchy estimator that uses a sign function as an instrument for the lagged level variable. They allow for strong cross-sectional dependence and time-varying variance. They show that the average of the individual specific IV-Cauchy test statistics converges to a standard normal distribution for panels where $N/T^{1/5} \to 0$. Westerlund (2014) proposes a Lagrange multiplier–type test for panels with heteroskedastic errors. The test procedure is brilliantly simple and yields a test statistic that has a noncentral χ^2 distribution, where the noncentrality parameter depends on the extent of cross-sectional dependence and heteroskedasticity.

4.1.4. Testing for stationarity. Most of the panel unit root tests have the null hypothesis of a unit root. However, they differ in the design of the alternative hypothesis. Consider the following two extreme alternatives: $\mathcal{H}_A^1$, where all time series in the panel are stationary, and $\mathcal{H}_A^2$, where a proportion of the time series in the panel are stationary. The first, $\mathcal{H}_A^1$, is said to be the homogeneous alternative. The second, $\mathcal{H}_A^2$, is said to be the heterogeneous alternative. In a short note, Pesaran (2012) comments on the interpretation of panel unit root tests. Rejecting the null of nonstationarity when the alternative hypothesis is $\mathcal{H}_A^1$ has a clear interpretation, that is, that all units are stationary. Two issues with this specification are that this test will have power even when not all of the units are stationary and that it can be taken as too restrictive, especially when the cross-section units have different dynamics. In contrast, $\mathcal{H}_A^2$ is less restrictive, but it is appropriate only when N is finite; otherwise, the test will lack power. Pesaran (2012) suggests that, for panels with large N and large T, it is more suitable to adopt an alternative hypothesis that lies in between $\mathcal{H}_A^1$ and $\mathcal{H}_A^2$. This can be stated as $\mathcal{H}_A^3$: $b(N)$ of the time series in the panel are stationary, where $b(N)$ is an increasing function of N. It is shown that the panel unit root tests have power if $\lim_{N\to\infty} \frac{b(N)}{N} = \delta$, where $0 < \delta \leq 1$. For instance, tests developed by Pesaran (2007b) and Pesaran et al. (2013) consider $\mathcal{H}_A^3$ as their alternative hypothesis. In this case, the interpretation of δ is straightforward. It gives the limiting proportion of the cross-section units that exhibit stationarity. This can be estimated consistently when the time series dimension is long enough. An estimator for δ is given by the number of units for which the null of nonstationarity is rejected by a univariate unit root test. The literature on estimating this δ, investigating the effects of δ on the cross-sectional variance, and identifying which units are stationary is still developing (Ng 2008, Pesaran 2007a, Smeekes 2015). Hanck (2013) proposes a test that allows the identification of the units in the panel for which the null of nonstationarity can be rejected.

4.1.5. Power of tests under local alternatives.

To investigate the behavior of test statistics under local alternatives, the autoregressive coefficient can be stated as $\theta_i = 1 - \frac{c_i}{f(N,T)}$, where c_i is random and $c_i \geq 0$ and $f(N, T)$ is a nondecreasing function of N and T. Then, under the assumption that c_i is independent and identically distributed across i with mean μ_c, the unit root testing problem can be stated as $\mathcal{H}_0 :\ \mu_c = 0$ against $\mathcal{H}_A^4 :\ \mu_c > 0$. This way of considering near unit root models enables the researchers to study the power of the tests for the neighborhoods of the nonstationary null (Moon & Perron 2004, Westerlund 2014, Westerlund et al. 2016).

4.1.6. Accounting for structural breaks.

Another issue that appears in unit root testing is the possible presence of structural breaks in the parameters of the models. Breaks can occur due to regime shifts, policy changes, or geological events. Not taking into account the structural changes might lead to misleading conclusions about the unit root properties of processes. Bai & Carrion-i-Silvestre (2009) show that the PANIC approach of defactoring the data is still valid in the presence of structural breaks in the mean of the series. If the structural break is in the slope of the time series, and the break point is not common for all cross-section units, then a modification of PANIC is required. They develop an iterative process to estimate the factors, break locations, and deterministic components. They then adopt a modified Sargan-Bhargava test to test for unit roots. Lee et al. (2016) modify the work of Pesaran et al. (2013) to allow for structural breaks.

4.2. Testing for Cointegration

Investigation of whether two or more variables share a common stochastic trend dates back to the 1980s. Since the seminal works of Engle & Granger (1987) and Johansen (1988), the analysis of long-run relationships has been a popular focus in econometrics. Two early literature reviews on cointegration analysis in panels were provided by Breitung & Pesaran (2008) and Choi (2013).

Suppose that the elements of the n_z-vector of panel variables $\mathbf{z}_{i,t} = (z_{1,i,t}, \ldots, z_{n_z,i,t})'$ for cross-section unit i at time t are all $I(1)$ variables, and suppose that, for each i,

$$\boldsymbol{\beta}_i'\mathbf{z}_{i,t} = \boldsymbol{\zeta}_{i,t} \sim I(0),$$

where $\boldsymbol{\beta}_i$ is a matrix of $m \times s$, and $\boldsymbol{\zeta}_{i,t}$ is an s—vector of $I(0)$ variables. We then say that the elements of $\mathbf{z}_{i,t}$ are cointegrated, and there are s linearly independent cointegrating relations between the elements of $\mathbf{z}_{i,t}$. One approach to testing for cointegration in this setup is to partition $\mathbf{z}_{i,t}$ into $z_{1,i,t}$ and $z_{2,i,t}, \ldots, z_{n_z,i,t}$, regress $z_{1,i,t}$ on $z_{2,i,t}, \ldots, z_{n_z,i,t}$, and check if the residuals from this regression are stationary. This type of test is called the residual-based test. These tests are appropriate when there is only one cointegrating relation, $s = 1$. When there is more than one cointegrating relation, $s > 1$, a system approach is more appropriate. When investigating the cointegration properties of panel data, one needs to take into account the potential cross-sectional dependence and heterogeneity and homogeneity concerns. For a list of works that assume cross-sectional independence in panel cointegration analysis, the reader is referred to the references of Breitung & Pesaran (2008). Violation of the cross-sectional independence assumption potentially invalidates the tests and leads to inefficiencies (Wagner & Hlouskova 2009). Some earlier works that allow for cross-sectional dependence are surveyed by Choi (2013).

The literature on residual-based tests addresses the issues of choice of defactorization, identification of the sources of nonstationarity—i.e., common stochastic trends or idiosyncratic stochastic trends—and identification of the cointegrating relationships. Gengenbach et al. (2006) and Bai & Carrion-i-Silvestre (2013) propose residual-based tests for no cointegration. Westerlund (2008) proposes a Hausman-type test to test for cointegration in panels. Westerlund & Edgerton (2007) propose a bootstrap to test for panel cointegration while allowing for cross-sectional dependence. Westerlund & Edgerton (2008) propose a simple test for no cointegration while allowing for heteroskedasticity and serial correlation in the errors, unit-specific time trends, cross-sectional dependence, and unknown structural breaks in the slope and intercept coefficient. Hanck (2009) combines the *p*-values of the time-series cointegration tests to obtain a panel cointegration test. This test is robust to cross-sectional dependence and heterogeneity. To achieve robustness against heterogeneity, Hanck (2009) uses a sieve bootstrap procedure. Chang & Nguyen (2012) use an IV approach to construct a test for cointegration in panels with endogeneity, cross-sectional dependence, and heterogeneity.

A system approach to a VAR framework is considered by Larsson & Lyhagen (2007). They adopt a likelihood-based framework and propose tests for the cointegrating rank of the system of homogeneous and heterogeneous long-run relations.

5. PANEL DATA MODELING ISSUES

5.1. Considerations When Estimating Common Factors

In the previous sections, we discuss estimation methods, statistical tests, and diagnostics for panel data that exhibit cross-sectional dependence. There are two strands of literature for dealing with cross-sectional dependence, namely, the spatial econometric approach and the residual-based multifactor approach. The spatial econometric approach relates the cross-sectional dependence on factors such as location and distance among panel units.

In the residual-based multifactor approach discussed by Chudik & Pesaran (2015b) and reviewed in this article, cross-sectional dependence between panel data units is characterized by a small number of unobserved common factors, possibly due to economy- or even worldwide shocks that affect all units, albeit with different intensities. Provided that they are identifiable and that one adopts the CCE approach proposed by Pesaran (2006) and presented in Section 3.1, these unobserved factors can be estimated by the cross-sectional averages of some observed time series or, alternatively, by the PC of the estimated covariance matrix of a set of variables that are closely linked to the unobserved factors. Both of these methods, the CCE- and the PC-based methods, rely on the use of point estimates or proxies of the unobserved factors and are therefore expected to be affected by estimation errors in small samples.

One could criticize the reluctance in this literature to postulate a joint model for the latent factors and for the panel data set, possibly a joint model conditional on a set of observed proxy variables. Such a joint model could be used to get optimal point estimates or proxies for the latent factors, given the joint postulated model. Optimal (efficient) estimates in a mean squared error sense would be obtained by using the conditional means of the factors for given data. This latter approach would also allow one to study and assess the properties of the proxies obtained by appropriate testing using the joint model. However, obviously, given that this estimator is a nonlinear function of the data, the required computations will be more involved, and the finite sample properties of the maximum likelihood estimator of the parameters and the factors might be inferior in small samples compared to those of more straightforward estimators of the factors, such as cross-sectional average- or PC-based estimators of the factors and parameters.

Alternatively, given a fully specified model for the latent factors, rather than using a conditional model for the data with the factors estimated by some proxies (substitution approach), one could obtain the marginal model for the data by integrating out the common factors (marginalization approach). To handle the latent factors, the substitution approach uses optimal estimates of the latent factors. The marginalization-based approach is a model-based approach that integrates out the latent factors. Provided that the model is well-specified, the marginalization-based approach is the preferred approach—at least asymptotically—as it is asymptotically efficient. However, in small samples and in situations with uncertainty about the process generating the latent variables, cross-sectional means, cross-sectional medians, and PC of variables correlated with the factors will indeed be sensible and robust alternatives to model-based estimates of the latent factors.

5.2. Parameter Plethora of Dynamic Panel Data Models

Economic panel data usually consist of multiple measurements of economic decisions or states of multiple economic units or individuals during multiple periods of time. While the amount of data is continuously increasing, the information available often does not meet the requirements for estimating and testing available econometric models. The time span of the observed data is sometimes too short to measure long-run effects with sufficient precision. Data are affected by latent individual effects and measurement errors, seasonal effects, etc. In practice, modeling the data means searching and testing for structure in the available data, designing proxies for missing variables, relying on other sources of information to design such proxies, etc. Information from sources such as the Internet could in some instances be useful when structured data are missing. Statistical procedures designed to search for structure in sets of so-called unstructured data and developments in the statistical learning theory are promising in this respect.

Designing specific models for specific types of data has often been the route to a promising outcome, as illustrated by the history of panel data econometrics. An example of such a specific model is the panel VAR model (see, e.g., Canova & Ciccarelli 2009, 2013) frequently used in macroeconomics and finance. Panel VARs have the same structure as VAR models. All variables are assumed to be endogenous and interdependent.

Let $\mathbf{Y}_t = (\mathbf{y}'_{1,t}, \mathbf{y}'_{2,t}, \ldots, \mathbf{y}'_{N,t})'$ be the stacked version of $\mathbf{y}'_{i,t}$, the row vector of G variables for each unit (country, sectors), $i = 1, \ldots, N$. A panel VAR is given by

$$\mathbf{y}_{i,t} = \mathbf{A}_{0,i}(t) + \mathbf{A}_i(L)\mathbf{Y}_{t-1} + \mathbf{F}_i(L)\mathbf{W}_t + \mathbf{u}_{i,t},$$

for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, where the $\mathbf{u}_{i,t}$ is a $G \times 1$ vector of random disturbances with covariance matrix $\mathbf{\Omega}_i$; the $\mathbf{u}_{i,t}$s are correlated across i (static interdependence); L denotes the lag operator; $\mathbf{W}_t$ denotes a vector of M exogenous or predetermined variables; and the coefficient matrices $\mathbf{A}_{0,i}(t)$, $\mathbf{A}_i(L)$, and $\mathbf{F}_i(L)$ could depend on the unit i. The vector $\mathbf{u}_t = (\mathbf{u}_{1,t}, \mathbf{u}_{2,t}, \ldots, \mathbf{u}_{N,t})'$ is independent and identically distributed $(\mathbf{0}, \mathbf{\Omega})$. Lags of all endogenous variables of all units enter the model for unit i (called dynamic interdependences).

The intercept, the slope, and the variances of the shocks $\mathbf{u}_{i,t}$ may be specific, a feature called cross-sectional heterogeneity by Canova & Ciccarelli (2013). They point out that a panel VAR has the same structure as large-scale VARs where dynamic and static interdependencies are allowed for by assumption. These features distinguish the panel VAR from VAR models used in micro studies.

Canova & Ciccarelli (2013) make a comparison of the main features of the panel VAR models with those of alternative models such as large-scale Bayesian VARs (e.g., Bańbura et al. 2010, De Mol et al. 2008), spatial econometric models (Anselin 2010), factor models (e.g., Stock & Watson 1989, 2005), global VARs (Dees et al. 2007, Pesaran et al. 2004), and bilateral panel VARs (e.g., Edelstein & Kilian 2009), and they compare and highlight similarities and differences among these models and panel VARs. In the analysis of a large-scale Bayesian VAR, the panel dimension of the data is not taken into consideration. Instead, all variables are treated symmetrically, and the Litterman-Minnesota-type prior (see Doan et al. 1984) that is used does not take cross-sectional information in the data into account. Spatial econometric and global VARs and factor models typically capture interdependencies within a set of factors or as a neighbor effect measured by physical distance.

5.3. Implications of Nonlinearity and Structural Changes for Panel Data Modeling

In recent years, the theory of statistical inference for nonlinear panel data has been developed.

Kapetanios et al. (2014) propose a nonlinear panel data model that generates both weak and strong cross-sectional dependence. A central assumption underlying the model is that agents' behavior is influenced by the views and actions of those around them. The model specification is flexible in terms of generating herding and clustering behavior.

While, when N is small, for instance when a macropanel for a small number of countries and a small set of variables is analyzed and the model is linear, rather than using panel data large N and large T asymptotics, one could analyze these data relying on finite sample properties or large T asymptotics for multivariate linear regressions (considering N to be fixed). When the model is nonlinear, one has to rely on large T asymptotics as finite sample properties of estimation methods are not known. However, when the model is nonlinear and the number of cross-sectional units is fairly large, panel data methods are required, so that it is important to study the behavior of methods for large T and reasonably large N cases, particularly when the model is nonlinear. For instance, Palm et al. (2012a,b) extend the work of Park & Phillips (1999, 2001) on nonlinear asymptotics for $I(1)$ regressors to nonlinear, nonstationary panel data models. They report results for the bias measured by the mean absolute error of the pooled nonlinear least squares estimator for independent cosummable, nonstationary panels with $I(1)$-variables and of the nonlinear least squares dummy variable estimator for independent nonstationary panels with $I(1)$ variables with fixed effects. The concept of cosummability of stochastic processes is introduced by Berenguer-Rico & Gonzalo (2014). Formally, the order of cosummability of a stochastic process gives a summary measure of the stochastic properties of persistence and evolution of the variance of the process without relying on a particular data generating process. Two processes y_t and x_t of order of integration of, respectively, δ_y and δ_x are cosummable if there exists $f(x_t)$ summable of order δ such that $u_t = y_t - f(x_t)$ is $S(\delta_u)$, with $\delta_u = \delta_y - \delta$ and $\delta > 0$.

The results obtained by Palm et al. (2012a,b) for nonlinear nonstationary panel data models indicate that asymptotic theory provides reliable guidance for these types of panels of N series, with N being between 10 and 30 and $T = 30$. These findings indicate that, for panel and time dimensions frequently given in empirical research in macroeconomics, we can expect to find reliable results.

If the nonlinear transformation is integrable-regular, then the rate of convergence of the pooled nonlinear least squares estimator for cross-sectionally independent homogeneous panels under cosummability is $T^{3/4}N^{1/2}$, whereas for homogeneous-regular nonlinear transformations, the rate of convergence is $TN^{1/2}$ when $T \to \infty$, followed by $N \to \infty$.

Thus, under nonstationarity, these rates are higher than those achieved for stationary variables. Cross-sectional independence and homogeneity often do not hold for economic panel data. If the time dimension is sufficiently large, with N being fixed (and not too large), as in multi-country macroeconomic forecasting and policy simulation, then one can also rely on large T-asymptotics (provided that T is sufficiently large) that have been obtained for (non)linear stationary and (non)stationary multivariate regressions and seemingly unrelated regressions.

In the econometric literature on VAR models, particularly in macroeconomic applications, most researchers advocate transforming economic variables to stationarity by log-differencing them or by including deterministic trend variables in the VAR model. The Wold or moving average representation (Wold 1938) or the impulse responses of this VAR model are obtained by inverting the VAR matrix. The Wold representation is uniquely determined up to the choice of the deterministic component and of one among the empirically equivalent triangular forms of the disturbance covariance matrix. The model is linear in the variables. Two of the concerns in empirical work are the choice of the transformation that makes the variables stationary and the choice of the number of lags in the VAR. Of course, as mentioned above, one can compare and test the VAR model against alternative nonlinear models to check whether there is nonlinearity present in the data that would not be detected by a (linear) VAR analysis. Such tests are also most appropriate if one has doubts about the appropriateness of the VAR model for the series at hand or in cases of structural breaks in the series, implying a change in the model structure. The innovations or impulses of the Wold representation should be serially uncorrelated, at least asymptotically, but not (necessarily) serially independent if the observed variables are second-order stationary and the correct order of the VAR is chosen. When (panel) data are stationary and generated by a nonlinear process with independent normally distributed disturbances, the innovations of the Wold representations of these variables will generally be neither normally distributed nor independent, although by construction, these innovations should be vector white noise. In the case where the generating process for the panel data is linear, and the data are jointly normally distributed, one can exactly pinpoint the statistical implications of this model for any Wold representation, both that of the joint process for all variables and that of the implied joint processes of subsets of the data from the panel under study.

A similar general statement can be made about modeling and testing for structural stability of panel data models. If a structural change occurred unexpectedly, its effect would materialize, at the earliest, at the moment of its occurrence. If the shape of the expected response in terms of changes in the structure of the panel data model is known a priori, then it could be modeled and translated into expected changes in the structure of the Wold representation used in the empirical analysis of the data. Given that the Wold representation of the panel model or submodel is nonlinear, in applied work, one likely has to carry out a detailed empirical analysis of the shape of the effects of a structural change in the data on the structure of the Wold representation used in the analysis.

5.4. Model and Parameter Uncertainty

A broad group of researchers in the field of panel data methods in economics and business supports the view that the dynamic models that have been developed and used should be interpreted as (usually linear) approximations of the conditional expectations of the variables modeled at period t given past values of these variables and, possibly, given a set of current and past values of some exogenous variables. If the variables involved are stationary, and no exogenous variables are included in the model, then these conditional expectations are in fact identical to the Wold representation of a stationary multivariate process referred to in Section 5.3. In this specific and often-used setting, model uncertainty basically entails uncertainty about the (weak) stationarity of the variables included in the model and about the chosen lag order of the Wold representation. Weak stationarity can be checked by various tests, for instance, by testing the lag order of the chosen Wold representation against that of higher lag order specifications, by comparing the linear Wold specification to nonlinear specifications, or by checking the stability over time of the Wold representation using structural stability tests. A point of concern might be the often implicitly assumed normality of the data, which is not implied by the Wold representation. As mentioned in Section 5.3, the innovation of Wold representations is white noise; that is, they are serially uncorrelated but not necessarily serially independent and normally distributed. In fact, if, to test the Wold specification, one had to make distributional assumptions, then it could be sensible to rely on fat-tail distributions and thereby explicitly account for the distributional uncertainty of the disturbances of the Wold representation.

Parameter uncertainty and model uncertainty in panel modeling have been effectively addressed by use of Bayesian methods. The successful adoption of a Bayesian approach in modeling and forecasting macroeconomic panel data using, among others, so-called Litterman's priors to address parameter uncertainty in Bayesian forecasting and policy analyses illustrates this (see, e.g., De Mol et al. 2008, Doan et al. 1984). To address model uncertainty, Bayesian methods have been applied to combine models and forecast using panels of international growth rates (see, e.g., Min & Zellner 1993).

6. ON THE USE OF PANEL DATA MODELS FOR PREDICTION AND POLICY MAKING

This section provides a review of research on macroeconomic forecasting using large panel factor models. D'Agostino & Giannone (2012) compare large panel data models to the static principal component approach of Stock & Watson (2002) and to the two-step approach of Forni et al. (2005) for forecasting industrial production and consumer price inflation. D'Agostino & Giannone (2012) conclude that both alternative approaches outperform the simple univariate autoregressive model. Also, few common factors are found to capture the predictable components of consumer price inflation and industrial production, and idiosyncratic dynamics appear to be negligible. The gain from factor-based predictions is substantial, and a few factors suffice.

Similar findings result from several empirical studies on macroeconomic forecasting undertaken by Arnold Zellner and various coauthors. For instance, Garcia-Ferrer et al. (1987) analyze the forecasting performance for GDP of nine countries and find that the use of world stock returns as an observed proxy of a common factor in a set of univariate autoregressive moving average models for the GDP of the nine countries significantly improved the forecasting performance. Furthermore, Zellner & Hong (1989) find that the forecast precision improved when the set of nine countries was extended by adding another nine countries and a world output growth rate as a second common factor. World output growth rate was measured by the cross-sectional median of the 18 countries' annual growth rates. The cross-sectional median was chosen as it is a more robust estimate of a common factor than the cross-sectional mean (advocated by H. Pesaran).

Finally, pooling with the objective of improving forecast precision in dynamic panel models in the presence of cross-sectional correlation that must be estimated may not always be the best thing to do, as has been shown by, e.g., Hoogstrate et al. (2000).

7. CONCLUDING REMARKS

This article provides an extensive review of the recent developments in the literature on cross-sectionally dependent panel data models with large time series and cross-section dimensions. We focus on the unobserved common factors approach to model cross-sectional dependence. We cover the topics that are necessary for a full-blown analysis of a panel data set. Our review includes guidelines for empirical research that uses panel data sets and critical discussions that can be of use for future theoretical research. As can be seen from our review, the literature on the analysis of large panel data models with cross-sectional dependence is considerably mature for standard models such as linear models with exogenous regressors, dynamic models, certain classes of nonstationary models, tests for unit roots, tests for cointegration, and tests for cross-sectional dependence. However, there are still questions for future work. These questions involve, for example, the issues that arise when dealing with unobserved common factors when estimating panel vector error correction models, testing for weak exogeneity, testing for cointegration rank using a systems approach, estimating vector threshold, using smooth transition models or models with other types of nonlinearities, and testing for unit roots when the alternative hypothesis contains processes with structural breaks. A second strand of future work includes expanding the applications of the methodologies to a wider set of fields, such as corporate finance, econometrics of climate change, predictability analysis, and cross-country growth analysis.

DISCLOSURE STATEMENT

Dr. Urbain did not get a chance to review this edited article before he passed away in October 2016. The other authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are grateful to our colleague Jean-Pierre Urbain who contributed to the completion of the first draft of this review. To our great regret, he passed away on October 1, 2016.

LITERATURE CITED

Amengual D, Watson MW. 2007. Consistent estimation of the number of dynamic factors in a large N and T panel. *J. Bus. Econ. Stat.* 25:91–96

Anselin L. 1988. *Spatial Econometrics: Methods and Models*, Vol. 4. Berlin: Springer

Anselin L. 2010. Thirty years of spatial econometrics. *Pap. Reg. Sci.* 89:3–25

Arellano M, Bond S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* 58:277–97

Arellano M, Bover O. 1995. Another look at the instrumental variable estimation of error-components models. *J. Econom.* 68:29–51

Bai J. 2003. Inferential theory for factor models of large dimensions. *Econometrica* 71:135–71

Bai J. 2009. Panel data models with interactive fixed effects. *Econometrica* 77:1229–79

Bai J, Carrion-i-Silvestre JL. 2009. Structural changes, common stochastic trends, and unit roots in panel data. *Rev. Econ. Stud.* 76:471–501

Bai J, Carrion-i-Silvestre JL. 2013. Testing panel cointegration with unobservable dynamic common factors that are correlated with the regressors. *Econom. J.* 16:222–49

Bai J, Kao C. 2006. On the estimation and inference of a panel cointegration model with cross-sectional dependence. In *Panel Data Econometrics: Theoretical Contributions and Empirical Applications*, ed. BH Baltagi, pp. 3–30. Bingley, UK: Emerald Group Publ.

Bai J, Kao C, Ng S. 2009. Panel cointegration with global stochastic trends. *J. Econom.* 149:82–99

Bai J, Ng S. 2002. Determining the number of factors in approximate factor models. *Econometrica* 70:191–221

Bai J, Ng S. 2004. A PANIC attack on unit roots and cointegration. *Econometrica* 72:1127–77

Bai J, Ng S. 2007. Determining the number of primitive shocks in factor models. *J. Bus. Econ. Stat.* 25:52–60

Bai J, Ng S. 2010. Panel unit root tests with cross-section dependence: a further investigation. *Econom. Theory* 26:1088–114

Bai J, Wang P. 2016. Econometric analysis of large factor models. *Annu. Rev. Econ.* 8:53–80

Bailey N, Kapetanios G, Pesaran MH. 2016. Exponent of cross-sectional dependence: estimation and inference. *J. Appl. Econom.* 31:929–60

Baltagi BH, Moscone F. 2010. Health care expenditure and income in the OECD reconsidered: evidence from panel data. *Econ. Model.* 27:804–11

Bańbura M, Giannone D, Reichlin L. 2010. Large Bayesian vector auto regressions. *J. Appl. Econom.* 25:71–92

Banerjee A, Marcellino M, Osbat C. 2004. Some cautions on the use of panel methods for integrated series of macroeconomic data. *Econom. J.* 7:322–40

Berenguer-Rico V, Gonzalo J. 2014. Summability of stochastic processes: a generalization of integration for non-linear processes. *J. Econom.* 178:331–41

Bertoli S, Moraga JF-H. 2013. Multilateral resistance to migration. *J. Dev. Econ.* 102:79–100

Blomquist J, Westerlund J. 2013. Testing slope homogeneity in large panels with serial correlation. *Econ. Lett.* 121:374–78

Blundell R, Bond S. 1998. Initial conditions and moment restrictions in dynamic panel data models. *J. Econom.* 87:115–43

Bond S, Leblebiciolu A, Schiantarelli F. 2010. Capital accumulation and growth: a new look at the empirical evidence. *J. Appl. Econom.* 25:1073–99

Breitung J, Das S. 2005. Panel unit root tests under cross-sectional dependence. *Stat. Neerlandica* 59:414–33

Breitung J, Pesaran MH. 2008. Unit roots and cointegration in panels. In *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, ed. L Matyas, P Sevestre, pp. 279–322. Berlin: Springer

Breusch TS, Pagan AR. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Rev. Econ. Stud.* 47:239–53

Bun MJ, Carree MA. 2005. Bias-corrected estimation in dynamic panel data models. *J. Bus. Econ. Stat.* 23:200–10

Canova F, Ciccarelli M. 2009. Estimating multicountry VAR models. *Int. Econ. Rev.* 50:929–59

Canova F, Ciccarelli M. 2013. Panel vector autoregressive models: a survey. In *VAR Models in Macroeconomics—New Developments and Applications: Essays in Honor of Christopher A. Sims*, ed. TB Fomby, L Kilian, A Murphy, pp. 205–46. Bingley, UK: Emerald Group Publ.

Cavaliere G. 2005. Unit root tests under time-varying variances. *Econom. Rev.* 23:259–92

Chamberlain G, Rothschild M. 1983. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* 51:1281–304

Chang Y, Nguyen CM. 2012. Residual based tests for cointegration in dependent panels. *J. Econom.* 167:504–20

Choi CY, Mark NC, Sul D. 2010. Bias reduction in dynamic panel data models by common recursive mean adjustment. *Oxf. Bull. Econ. Stat.* 72:567–99

Choi I. 2013. *Panel cointegration*. Work. Pap. 1208, Res. Inst. Market Econ., Sogang Univ., Seoul

Chudik A, Pesaran MH. 2011. Infinite-dimensional VARs and factor models. *J. Econom.* 163:4–22

Chudik A, Pesaran MH. 2013. Econometric analysis of high dimensional VARs featuring a dominant unit. *Econom. Rev.* 32:592–649

Chudik A, Pesaran MH. 2015a. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *J. Econom.* 188:393–420

Chudik A, Pesaran MH. 2015b. Large panel data models with cross-sectional dependence: a survey. In *The Oxford Handbook of Panel Data*, ed. BH Baltagi, pp. 3–45. Oxford, UK: Oxford Univ. Press

Chudik A, Pesaran MH, Tosetti E. 2011. Weak and strong cross-section dependence and estimation of large panels. *Econom. J.* 14:C45–90

Coakley J, Fuertes AM, Smith R. 2002. *A principal components approach to cross-section dependence in panels*. Unpublished manuscript, Birkbeck Coll., Univ. London

Coakley J, Fuertes AM, Smith R. 2006. Unobserved heterogeneity in panel time series models. *Comput. Stat. Data Anal.* 50:2361–80

D'Agostino A, Giannone D. 2012. Comparing alternative predictors based on large-panel factor models. *Oxf. Bull. Econ. Stat.* 74:306–26

De Mol C, Giannone D, Reichlin L. 2008. Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *J. Econom.* 146:318–28

Dees S, di Mauro F, Pesaran MH, Smith LV. 2007. Exploring the international linkages of the euro area: a global VAR analysis. *J. Appl. Econom.* 22:1–38

Demetrescu M, Hanck C. 2012. Unit root testing in heteroscedastic panels using the Cauchy estimator. *J. Bus. Econ. Stat.* 30:256–64

Dickey DA, Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* 74:427–31

Doan T, Litterman R, Sims C. 1984. Forecasting and conditional projection using realistic prior distributions. *Econom. Rev.* 3:1–100

Eberhardt M, Helmers C, Strauss H. 2013. Do spillovers matter when estimating private returns to R&D? *Rev. Econ. Stat.* 95:436–48

Eberhardt M, Teal F. 2011. Econometrics for grumblers: a new look at the literature on cross-country growth empirics. *J. Econ. Surv.* 25:109–55

Eberhardt M, Teal F. 2013. No mangoes in the tundra: spatial heterogeneity in agricultural productivity analysis. *Oxf. Bull. Econ. Stat.* 75:914–39

Edelstein P, Kilian L. 2009. How sensitive are consumer expenditures to retail energy prices? *J. Monet. Econ.* 56:766–79

Engle RF, Granger CW. 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55:251–76

Ergemen YE, Velasco C. 2017. Estimation of fractionally integrated panels with fixed effects and cross-section dependence. *J. Econom.* 196:248–58

Ertur C, Musolesi A. 2017. Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion. *J. Appl. Econom.* 32:477–503

Everaert G, De Groote T. 2016. Common correlated effects estimation of dynamic panels with cross-sectional dependence. *Econom. Rev.* 35:428–63

Forni M, Hallin M, Lippi M, Reichlin L. 2005. The generalized dynamic factor model: one-sided estimation and forecasting. *J. Am. Stat. Assoc.* 100:830–40

Garcia-Ferrer A, Highfield RA, Palm F, Zellner A. 1987. Macroeconomic forecasting using pooled international data. *J. Bus. Econ. Stat.* 5:53–67

Gengenbach C, Palm FC, Urbain JP. 2006. Cointegration testing in panels with common factors. *Oxf. Bull. Econ. Stat.* 68:683–719

Gengenbach C, Palm FC, Urbain JP. 2009. Panel unit root tests in the presence of cross-sectional dependencies: comparison and implications for modelling. *Econom. Rev.* 29:111–45

Hahn J, Kuersteiner G. 2002. Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. *Econometrica* 70:1639–57

Hanck C. 2009. Cross-sectional correlation robust tests for panel cointegration. *J. Appl. Stat.* 36:817–33

Hanck C. 2013. An intersection test for panel unit roots. *Econom. Rev.* 32:183–203

Harding M, Lamarche C. 2011. Least squares estimation of a panel data model with multifactor error structure and endogenous covariates. *Econ. Lett.* 111:197–99

Holly S, Pesaran MH, Yamagata T. 2010. A spatio-temporal model of house prices in the USA. *J. Econom.* 158:160–73

Hoogstrate AJ, Palm FC, Pfann GA. 2000. Pooling in dynamic panel-data models: an application to forecasting GDP growth rates. *J. Bus. Econ. Stat.* 18:274–83

Hsiao C, Pesaran MH, Tahmiscioglu AK. 2002. Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *J. Econom.* 109:107–50

Hurlin C, Mignon V. 2007. *Second generation panel unit root tests*. Work. Pap., HAL, Cent. Commun. Sci. Dir., Lyon, France

Johansen S. 1988. Statistical analysis of cointegration vectors. *J. Econ. Dyn. Control* 12:231–54

Juodis A, Karabiyik H, Westerlund J. 2018. *On the robustness of the pooled CCE estimators*. Work. Pap., Vrije Univ. Amsterdam

Kao C, Trapani L, Urga G. 2012. Asymptotics for panel models with common shocks. *Econom. Rev.* 31:390–439

Kapetanios G, Mitchell J, Shin Y. 2014. A nonlinear panel data model of cross-sectional dependence. *J. Econom.* 179:134–57

Kapetanios G, Pesaran MH, Yamagata T. 2011. Panels with non-stationary multifactor error structures. *J. Econom.* 160:326–48

Karabiyik H, Reese S, Westerlund J. 2017. On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *J. Econom.* 197:60–64

Karabiyik H, Urbain JP, Westerlund J. 2019. CCE estimation of factor-augmented regression models with more factors than observables. *J. Appl. Econom.* 34(2):268–84

Kiviet JF. 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *J. Econom.* 68:53–78

Lancaster T. 2002. Orthogonal parameters and panel data. *Rev. Econ. Stud.* 69:647–66

Larsson R, Lyhagen J. 2007. Inference in panel cointegration models with long panels. *J. Bus. Econ. Stat.* 25:473–83

Lee C, Wu JL, Yang L. 2016. A simple panel unit-root test with smooth breaks in the presence of a multifactor error structure. *Oxf. Bull. Econ. Stat.* 78:365–93

Lee L-F, Yu J. 2010. Some recent developments in spatial panel data models. *Reg. Sci. Urban Econ.* 40:255–71

Levin A, Lin C-F, Chu C-SJ. 2002. Unit root tests in panel data: asymptotic and finite-sample properties. *J. Econom.* 108:1–24

Min C, Zellner A. 1993. Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *J. Econom.* 56:89–118

Moon HR, Perron B. 2004. Testing for a unit root in panels with dynamic factors. *J. Econom.* 122:81–126

Moon HR, Phillips PC. 1999. Maximum likelihood estimation in panels with incidental trends. *Oxf. Bull. Econ. Stat.* 61:711–47

Moon HR, Weidner M. 2015. Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica* 83:1543–79

Moon HR, Weidner M. 2017. Dynamic linear panel regression models with interactive fixed effects. *Econom. Theory* 33:158–95

Moreira MJ. 2009. A maximum likelihood method for the incidental parameter problem. *Ann. Stat.* 37:3660–96

Ng S. 2008. A simple test for nonstationarity in mixed panels. *J. Bus. Econ. Stat.* 26:113–27

Nickell S. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49:1417–26

O'Connell PG. 1998. The overvaluation of purchasing power parity. *J. Int. Econ.* 44:1–19

Palm F, Urbain JP, Wan L. 2012a. *Asymptotic theory of pooled nonlinear least squares estimation for panel co-summable regressions*. Work. Pap., Grad. School Bus. Econ., Maastricht Univ., Neth.

Palm F, Urbain JP, Wan L. 2012b. *Nonlinear least squares dummy variables estimation for independent co-summable panels with fixed effects*. Work. Pap., Grad. School Bus. Econ., Maastricht Univ., Neth.

Park JY, Phillips PC. 1999. Asymptotics for nonlinear transformations of integrated time series. *Econom. Theory* 15:269–98

Park JY, Phillips PC. 2001. Nonlinear regressions with integrated time series. *Econometrica* 69:117–61

Pesaran MH. 2004. *General diagnostic tests for cross section dependence in panels*. Work. Pap. 1229, CESifo, Munich, Ger.

Pesaran MH. 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74:967–1012

Pesaran MH. 2007a. A pair-wise approach to testing for output and growth convergence. *J. Econom.* 138:312–55

Pesaran MH. 2007b. A simple panel unit root test in the presence of cross-section dependence. *J. Appl. Econom.* 22:265–312

Pesaran MH. 2012. On the interpretation of panel unit root tests. *Econ. Lett.* 116:545–46

Pesaran MH. 2015. Testing weak cross-sectional dependence in large panels. *Econom. Rev.* 34:1089–117

Pesaran MH, Kapetanios G. 2005. *Alternative approaches to estimation and inference in large multifactor panels: small sample results with an application to modelling of asset returns*. Work. Pap. 1416, CESifo, Munich, Ger.

Pesaran MH, Schuermann T, Weiner SM. 2004. Modeling regional interdependencies using a global error-correcting macroeconometric model. *J. Bus. Econ. Stat.* 22:129–62

Pesaran MH, Smith LV, Yamagata T. 2013. Panel unit root tests in the presence of a multifactor error structure. *J. Econom.* 175:94–115
Pesaran MH, Tosetti E. 2011. Large panels with common factors and spatial correlation. *J. Econom.* 161:182–202
Pesaran MH, Ullah A, Yamagata T. 2008. A bias-adjusted LM test of error cross-section independence. *Econom. J.* 11:105–27
Pesaran MH, Yamagata T. 2008. Testing slope homogeneity in large panels. *J. Econom.* 142:50–93
Phillips PC, Sul D. 2003. Dynamic panel estimation and homogeneity testing under cross section dependence. *Econom. J.* 6:217–59
Phillips PC, Sul D. 2007. Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *J. Econom.* 137:162–88
Reese S, Westerlund J. 2016. Panicca: panic on cross-section averages. *J. Appl. Econom.* 31:961–81
Ross S. 1976. Arbitrage theory of capital asset pricing. *J. Econ. Theory* 13:341–60
Sarafidis V, Robertson D. 2009. On the impact of error cross-sectional dependence in short dynamic panel estimation. *Econom. J.* 12:62–81
Sarafidis V, Wansbeek T. 2012. Cross-sectional dependence in panel data analysis. *Econom. Rev.* 31:483–531
Sarafidis V, Yamagata T, Robertson D. 2009. A test of cross section dependence for a linear dynamic panel model with regressors. *J. Econom.* 148:149–61
Smeekes S. 2015. Bootstrap sequential tests to determine the order of integration of individual units in a time series panel. *J. Time Ser. Anal.* 36:398–415
Song M. 2013. *Asymptotic theory for dynamic heterogeneous panels with cross-sectional dependence and its applications*. Unpublished manuscript, Dep. Econ., Columbia Univ., New York
Stock JH, Watson MW. 1989. New indexes of coincident and leading economic indicators. *NBER Macroecon. Annu.* 4:351–94
Stock JH, Watson MW. 2002. Forecasting using principal components from a large number of predictors. *J. Am. Stat. Assoc.* 97:1167–79
Stock JH, Watson MW. 2005. Understanding changes in international business cycle dynamics. *J. Eur. Econ. Assoc.* 3:968–1006
Teles VK, Mussolini CC. 2014. Public debt and the limits of fiscal policy to increase economic growth. *Eur. Econ. Rev.* 66:1–15
Wagner M, Hlouskova J. 2009. The performance of panel cointegration methods: results from a large scale simulation study. *Econom. Rev.* 29:182–223
Westerlund J. 2008. Panel cointegration tests of the Fisher effect. *J. Appl. Econom.* 23:193–233
Westerlund J. 2014. Heteroscedasticity robust panel unit root tests. *J. Bus. Econ. Stat.* 32:112–35
Westerlund J. 2015a. The effect of recursive detrending on panel unit root tests. *J. Econom.* 185:453–67
Westerlund J. 2015b. The power of panic. *J. Econom.* 185:495–509
Westerlund J. 2015c. Rethinking the univariate approach to panel unit root testing: using covariates to resolve the incidental trend problem. *J. Bus. Econ. Stat.* 33:430–43
Westerlund J. 2018. CCE in panels with general unknown factors. *Econom. J.* 21:264–76
Westerlund J, Edgerton DL. 2007. A panel bootstrap cointegration test. *Econ. Lett.* 97:185–90
Westerlund J, Edgerton DL. 2008. A simple test for cointegration in dependent panels with structural breaks. *Oxf. Bull. Econ. Stat.* 70:665–704
Westerlund J, Hosseinkouchack M. 2016. Modified CADF and CIPS panel unit root statistics with standard chi-squared and normal limiting distributions. *Oxf. Bull. Econ. Stat.* 78:347–64
Westerlund J, Hosseinkouchack M, Solberger M. 2016. The local power of the CADF and CIPS panel unit root tests. *Econom. Rev.* 35:845–70
Westerlund J, Urbain JP. 2013. On the estimation and inference in factor-augmented panel regressions with correlated loadings. *Econ. Lett.* 119:247–50
Westerlund J, Urbain JP. 2015. Cross-sectional averages versus principal components. *J. Econom.* 185:372–77
Wold H. 1938. *A Study in the Analysis of Stationary Time Series*. Uppsala, Swed.: Almqvist & Wiksell
Zellner A, Hong C. 1989. Forecasting international growth rates using Bayesian shrinkage and other procedures. *J. Econom.* 40:183–202

Annual Review of Economics

Volume 11, 2019

Contents

Symposium: Universal Basic Income

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://www.annualreviews.org/errata/economics